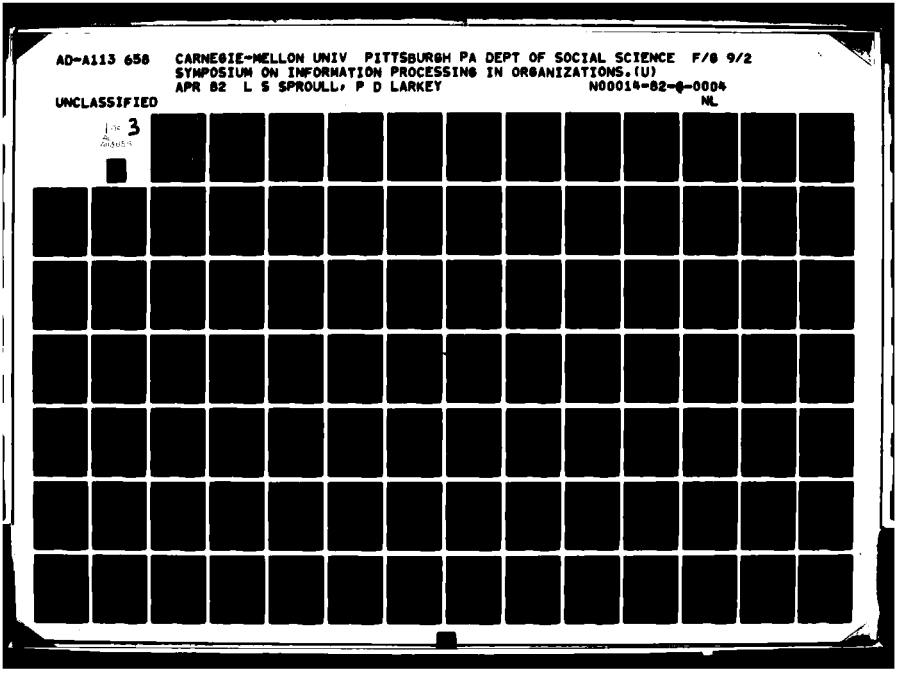
AD-A113 658 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF SOCIAL SCIENCE F/G 9/2
SYMPOSIUM ON INFORMATION PROCESSING IN ORGANIZATIONS.(U)
APR 82 L S SPROULL, P D LARKEY N00014-82-G-0004
UNCLASSIFIED NL



SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. AD A113658
3. RECIPIENT'S CATALOG NUMBER	
TLE (and Subtitle) SYMPOSIUM ON INFORMATION PROCESSING IN ORGANIZATIONS	5. TYPE OF REPORT & PERIOD COVERED October 81 - April 82 Final Report -
	6. PERFORMING ORG. REPORT NUMBER
JTHOR(s) Lee S. Sproull Patrick D. Larkey	8. CONTRACT OR GRANT NUMBER(s) N00014-82-G-0004
ERFORMING ORGANIZATION NAME AND ADDRESS Department of Social Science Carnegie-Mellon University Pittsburgh, PA. 15213	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ONR 170-936
CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (442) Organizational effectiveness group Arlington, VA 22217	12. REPORT DATE April 82
	13. NUMBER OF PAGES
MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Organizational information processing
Organizational decision making

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A two-day symposium was held at Carnegie-Mellon University in October 1981 to bring together social scientists doing research in the area of information processing in organizations. Eight papers were presented and discussed in detail. Promising areas for future research were identified.

include :

DTIC ELECTE APR 21 1982 C H

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DTIC FILE COPY

82 04 12 140

SYMPOSIUM ON INFORMATION PROCESSING IN ORGANIZATIONS

Final Technical Report

Lee S. Sproull
Patrick D. Larkey
Carnegie-Mellon University

On October 16-17, 1982, a Symposium on Information Processing in Organizations was held at Carnegie-Mellon University. Eight papers were delivered by social scientists. The authors, affiliations and paper titles are:

Richard Cyert and Morris DeGroot, C-MU: "The ~~minimization~~ maximization process under uncertainty"

Peter Keen, MIT: "Information systems in organizations"

Patrick Larkey and Richard Smith, C-MU: "Formulating and justifying budget problems"

Elizabeth Loftus, University of Washington: "Communicating with people in emergencies"

James March, Stanford University: "Gossip, information and decision making"

Allen Newell, C-MU: "An on-going case study in technological innovation"

J. Rounds, UCLA: "Information and ambiguity in organizational change"

Lee Sproull, C-MU: "The nature of managerial attention"

Discussions were led by:

Garry Brewer, Yale University

John P. Crecine, C-MU

Herbert Simon, C-MU

Joel Tarr, C-MU

Copies of the papers are attached.

This paper provides background material for Al Newell's presentation.

ZOG and the USS CARL VINSON

Allen Newell, Donald McCracken, George Robertson, Robert Akscyn
Computer Science Department
Carnegie-Mellon University

To appear in the Computer Science Research Review 1980 - 1981

17 September 1981

DRAFT

The ZOG research at CMU is supported by the Office of Naval Research under contract N00014-76-0874, and also partially by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 3597, monitored by the Air Force Avionics Laboratory under contract F33615-78-C-1551.

Table of Contents

List of Figures

ZOG and the USS CARL VINSON

In the standard view, the application of science goes through four stages: *research, development, prototype and production*. Each stage is successively more directed, involves larger investments of capital and takes different types of people. Yet, in fact, each scientific field evolves its own unique patterns for application. Computer science is still in a fluid state in this respect--indeed, it may never become tranquil enough to solidify a permanent pattern. However, the indications are that, compared to the standard picture, the path to application is often more compressed, more intricately interwoven, and involves more of the same people throughout.

The ZOG/VINSON project provides an interesting case study in how a direct application of research can happen in computer science. Though only in mid course--and far from being assured of success--the project is far enough along to be worth discussing. First, we will describe ZOG, which is an experimental human-computer interface and stands at the center of the whole project. Then we will give a brief history of the ZOG project at CMU and how we became involved with developing a ZOG system on the USS CARL VINSON, the newest nuclear powered aircraft carrier, now being outfitted at Newport News, Virginia. Next we will describe the scope of the project. This will provide enough background to reflect on the issue raised in the introductory paragraph. In particular, what *makes possible direct application of computer science research* and, second, what makes such an application interesting enough for computer scientists to become involved.

1. What is ZOG?

Put in a single sentence, ZOG is a rapid-response large-network menu-selection computer interface [11]. The user communicates to the computer by selecting from a menu of displayed and explained choices. The computer communicates back to the user by displaying some information along with an additional menu of selections (it may also execute programs). Menu selection is a common mode of user-computer communication. What distinguishes ZOG is its full exploitation of this mode -- the unlimited use of displays and selections. This implies the existence of a large network of display *frames* (as they are called) that give information and permit selection of yet other frames. These *ZOGnets* can be tens of thousands of frames. The strength of menu selection is its self-explanatory character; its weakness is that, as expertise and familiarity increase, selecting through a tree of options appears clumsy and time consuming. Thus, the final key aspect of ZOG is that the selection is very fast, independent of which frame in the entire network is being accessed.

Figure 1-1 shows a typical ZOG frame. Working down from the top, there is a *title*, followed by a

body of *text* which displays the information offered by the frame. Next is the menu of *options*, each distinguished by a single character (1 to 3 in the figure). Each option can link to another frame, and typing the assigned character replaces the current frame with this next frame. The *option text* associated with each selection explains what the user will find. There are other menu selections on the frame, called *pads*. The row at the bottom (the *global pads*), is available on every frame and provides a basic set of actions and links for searching the ZOGnet, entering the editor, going to a root frame, etc. The others (the *local pads*) provide similar functions, but specialized to a particular subnetwork. As the pads indicate, selections can execute programs as well as link to a new frame.

```
This TITLE line summarizes the frame's content                    Frame1

This TEXT expands the frame's main point of information. It is often omitted.
The options below can provide an enumerated expansion.

  1. This OPTION leads to another frame

  2. OPTIONs often are like subpoints in an outline

  3.-The minus sign means this OPTION has no next frame


                L. This LOCAL PAD is a cross-reference link

                A. Local pads can also execute actions




edit  help  back next mark return  zog  display user top shell goto find  info
```

Figure 1-1: A Self-Describing, Typical ZOG Frame

To be concrete, suppose the user was faced with the frame in Figure 1-1, which was part of a ZOGnet describing the ZOG system. Wanting to know more, the user might select option 2 (by typing the character 2 on the keyboard). Immediately, the frame would be replaced by the frame shown in Figure 1,2. The user might want to go on and select another frame (e.g., at option 1 of Figure 1-2) to get more information. Alternatively, he might want to go back to the frame of Figure 1-1, which he could do by selecting the global pad, back (typing character b), or to return to the top of the net, by selecting top (typing t), etc.

With every frame leading directly to a dozen other frames, an entire network is created. Figure 1-3 tries to give some idea of this, though it can't show all the links from global pads back to the root frame (zog) or other top frames. It does show a user-generated path though the ZOGnet (the dashed

```
OPTIONS are often like subpoints in an outline                    Frame6

This TEXT expands on the point made above, if necessary.
Note that in keeping with ZOG convention, the TITLE above is the same as the
text of the OPTION that led here from the previous frame.

  1. This OPTION gives yet further information about some aspect of the
     main point of this frame.

  2.-This OPTION makes a second subpoint, but leads no further.


            L. LOCAL PAD

            A. Active LOCAL PAD




edit  help  back next mark return  zog  display user top shell goto find  info
```

Figure 1-2: Another Example ZOG Frame

line). The back pad permits retreat back up this path step by step. The return global pad presents all of the frames of this pad as a menu of options, so the user can jump back to any prior point along the path. ZOGnets for little tasks run 50-100 frames; total ZOGnets to do a large job run into the thousands or tens of thousands. Although Figure 1-3 implies a certain locality in the ZOGnet, the time from one frame to another arbitrarily located frame is designed to be constant.

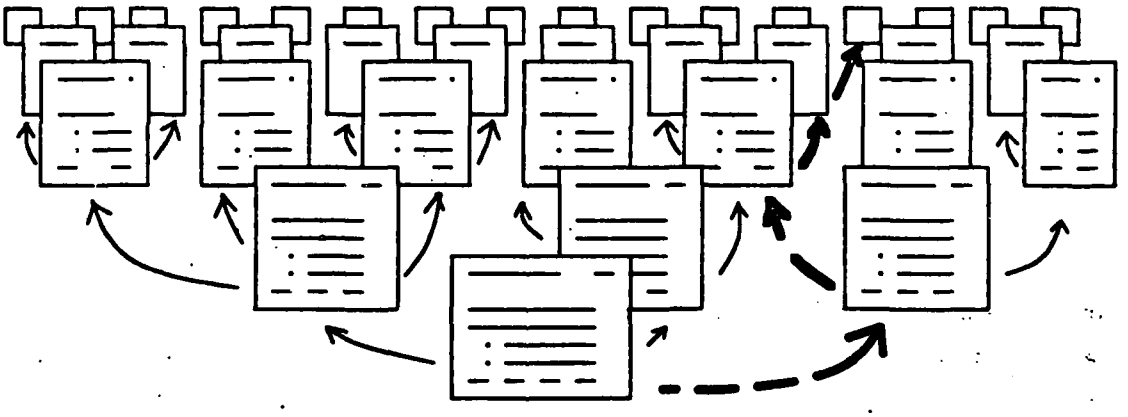


Figure 1-3: A ZOGnet

ZOG has other features beside the basic search and explore capability just illustrated. One is an editor which permits general editing of a frame, the creation of new frames and the linking of one frame to another. This editor (called ZED) is available at every frame via the global pad edit. ZOG is thus meant to be a dynamic memory structure, not just a retrieval system. A second feature is that ZOG is a communications multiplexer, which sits between the user and a general computer. It can modify messages between the user and other programs running on the computer. An important result is that a ZOG application system consists not only of a ZOGnet, but also of a collection of *agents*. These are programs evoked by the user through the ZOGnet, which can manipulate the ZOGnet and do other jobs. For instance, agents exist that create a document for a subnet in an appropriate and pleasing external form. They do this by working through *Scribe*, a modern document production system [9]. (The present paper was created in ZOG and exists both as hard copy and as a ZOGnet.) Besides the editor and the agent capability, there is a general search capability for character strings and a statistical capability for studying the behavior of the user.

ZOG is a general purpose interface, though especially suited to situations with extensive communication to the user of data that cannot easily be algorithmically generated. This fits a wide range of tasks. One is a *data base* where browsing is important, e.g., the characterization of the literature in a field, or a browsing system for a library [2]. Another is a *guidance system*, where the ZOGnet holds instructions about how to use a complex program. This was the original 1972 use, described in the following section; but more recently ZOG has been used to provide a guidance system for an elaborate AI medical diagnosis system, called INTERNIST [7]. *Management nets* constitute another type of use, where a project keeps all of its records, communication, discussion, etc., in a ZOGnet and expands it freely, under the impact of the project members. We have used such a ZOGnet for ZOG itself, and the ZOG/Vinson application has many similar features.

2. History

ZOG's beginnings go back to 1972, when a summer workshop was held for a group of cognitive psychologists interested in simulation [6]. One problem was to provide a way for the group to get hands-on experience with several large simulation programs developed by independent investigators, without knowing anything about the programs themselves or the operating system of the computer on which they were to run. Three of us (GR, DMC, AN) developed a program to act as a uniform interface for gaining access to all these programs, providing guidance in their use. This interface was called ZOG (now we call it ZOG-1) and had most of the characteristics of the system described here. Its critical deficiency was the terminal technology, which was only 300 baud hardcopy. Though ZOG-1 worked, the response rate was sufficiently mismatched to the requirements of the interface that

further development was not warranted, and we turned our attention to other things.

However, in 1975 we came into contact with PROMIS (Problem Oriented Medical Information System), an effort by Dr. Lawrence Weed of the University of Vermont Medical School to provide a total information system to support his notion of a problem-oriented approach [13] [15]. Two of us (AN, GR) served on a technical advisory committee for PROMIS. PROMIS turned out to be a system remarkably like ZOG, but carried much further. It used a terminal technology that was up to the demands for interaction, namely, a touch-screen system that provided about 1/2 second response for a frame. They had built a network of some 35,000 frames and demonstrated that a total complex task could be handled with this system philosophy. PROMIS rekindled our interest in such systems, for it showed the interface to be not only viable, but quite unique in its characteristics.

PROMIS was dedicated to its medical application, and we decided that this type of interface should be investigated in a larger computer science context. Thus, with ONR support, we began a small effort in 1975 to create a new version of ZOG, paying close attention to the PROMIS experience, that would live in a general interactive environment. An important requirement was that ZOG should be a general agent through which arbitrary other, pre-existing programs could be executed. Our primary objective was to assess whether the interface was genuinely interesting.

In pursuit of this goal we have brought up several versions of ZOG on the PDP10, under both TOPS10 and TOPS20 monitors, on C.mmp, the department's experimental multiprocessor, and most recently on the VAX/11-780. ZOG has been coded in our own implementation language, L* [5], which has made it easy to transfer across machines. We used ZOG for a number of explorations: as guidance systems, as issue nets (holding the growing data base of discussion on a complex issue), as literature data bases, and as library browsing catalogues. The main impediments were hardware and operating-system constraints. First, the Computer Science Department computing environment was interactive, but at 1200 baud, which turned out to be barely acceptable and far below appropriate rapid response. Second, the operating system imposed delays on the execution of subjobs (i.e., jobs executed through ZOG) that inhibited extensive use. The C.mmp experience is instructive here. We used C.mmp in order to obtain a terminal with instant response (less than 0.1 sec). But with only two such special terminals, the basic mode of operation remained at the community level of 1200 baud.

It is useful to understand the evolution of our computing environment. Throughout the seventies it has been built around multiple time-shared PDP10s with, as noted, 1200 baud access. Starting in 1980 we began to acquire DEC VAXs. We now have a number of these connected together (with the 10s) through an Ethernet local network, which also includes a number of Xerox Altos. Because of the

load on the 10s, ZOG was moved to the VAX very early. In parallel with the acquisition of the VAXs the department planned a five year effort to create a scientific personal integrated computing environment (SPICE) [4] [3]. The Three Rivers PERQ was selected as a pre-SPICE personal computer, in order to design and bring up an initial basic system [12]. There are close ties between ZOG and the SPICE effort, in part because of shared workers, in part because SPICE is a cradle within which many software efforts in the environment will eventually integrate.

In February 1980, Capt. Richard Martin visited the project. From our point of view he was simply another visitor. From the Captain's point of view the visit was not at all casual. He was assuming command of the new nuclear powered aircraft carrier, USS CARL VINSON, just then being launched. His responsibility was to outfit the carrier and give her an initial character. He had decided to incorporate whatever he could of the advanced research in computer science and artificial intelligence, making the carrier a test bed for getting scientific advances into the service of the fleet. He was visiting many of the ONR-supported research sites, to understand the current state of the art and to assess whether anything was at an appropriate state of development. Since he was temporarily in Pittsburgh at Westinghouse Bettis Atomic Power Lab, an intense interaction ensued about ZOG and its characteristics.

One of the Captain's goals was a thorough overhaul of the way the Carl Vinson would handle its administrative operations. He became convinced that a ZOG-like system, implemented on a modern locally networked collection of powerful work stations, would be feasible and would move him a long ways toward his goal of the test bed. We were, to say the least, dubious. Although we were looking for an appropriate driving application, we were hardly in the market for an application with the large scale, real time, real consequence characteristics of the system that the Captain was proposing. However, the Captain had strong arguments for the payoff to the Navy from ONR-supported research, and good answers for most of our objections. We agreed to proceed, after laying down some important constraints (to be mentioned later) to help assure the chances of success. The project actually got under way in July 1980, only a few months after the Captain had first made contact.

3. The project

We now have enough background to describe the project as it currently stands. The objective is to establish on the USS CARL VINSON a system capable of supporting certain management functions. As Figure 3-1 shows, the total system has a substantial number of levels. Working up from the bottom, there is the hardware system (PERQs on an Ethernet network, for the Vinson system). On top of that comes an operating system (SALT), and a programming language system (Pascal). Next come ZOG

and Scribe, which can be likened to application-oriented higher-level languages. Then comes the specific system (ZOGnet format design plus agents) for performing the particular management functions; this is analogous to an application program. Above that comes the actual large ZOGnet, analogous to the content of a data base. Finally, at the top, there is a level dealing with the entire range of system maintenance functions: installing the system, training the ship's personnel in its use, making the system work, modifying it as experience accumulates, and gradually enhancing its functions. Note that the CMU and CARL VINSON systems share everything at the Application Level and above (plus substantial parts of the Software Level), which of course is the point of using the CMU system as a development base.

LEVEL	CMU SYSTEM	CARL VINSON SYSTEM
Maintenance	Facilities for installation, training, operation, maintenance, and evolution	
Content	------ Contents of the ZOGnet ------ (functionally complete subset)	(continued growth)
Application	------ ZOGnet formats, Agents ------ (development, experimental use)	(iteration, operational use)
Software	VAX ZOG, Scribe	PERQ ZOG, Scribe
Language	L*, Pascal, C	Pascal
Operating System	Unix	SALT
Hardware	VAX	PERQ, Ethernet

Figure 3-1: Levels of the System

3.1. The management functions

With the entire structure laid out in Figure 3-1, we can start with the management tasks to be done. A modern carrier, carrying a total complement of 5600 men and living at sea for long periods of time, has the full range of management tasks and problems [14]. Two specific management tasks have been isolated out for attention in the ZOG/VINSON project: The maintenance and use of the Ship's Organization and Regulations Manual (SORM); and high level Planning and Evaluation (P&E).

The SORM is the repository of much of the standard procedures for a ship's operation. It describes what to do to accomplish all the things that have to be done. It is used in conjunction with directives that come down from higher commands. Much of the information in the SORM also exists in the

heads of experienced crew members, but the SORM still plays several critical roles. It is the court of last resort, the source of continuity, and the means of training new crew members. SORMs are large documents (the one for the USS EISENHOWER is 357 pages) which are continually changing. They are of varying degrees of effectiveness. One of the Captain's goals for the CARL VINSON is to produce a SORM which covers the totality of actual operations and which is highly effective. Putting the SORM into ZOG as a large ZOGnet (estimated 20,000 frames, with much more detail than the EISENHOWER SORM) is the chief means to this goal. The SORM for a major ship, such as the CARL VINSON, does not pre-exist. Thus, creating the SORM as a structured ZOGnet acts as a strong force to rationalize it and analyze all of the operations.

Planning occurs quite generally at all levels of operation. In large organizations some formal mechanisms are required, to effect coordination and to permit evaluation. The second management function to be provided by ZOG is support for on-line planning. Plans will exist in an integrated ZOGnet, which will be updated and modified continually as each plan is extended and changed. Exploration of plans from different perspectives will be possible, e.g., by task, by persons or by resources. Some automatic monitoring of plans for consistency and critical events, and some propagation of status through plans will be possible. The ZOG P&E will be aimed at the top levels of shipboard operation, down to the level of the department. There will be one or more PERQ work stations in each of the 17 departments. However, there are not enough work stations for adequate on-line operation at the next level of organization, the division, which has some 70 units.

Maintaining the SORM and planning are independent management functions to be performed on the USS CARL VINSON, hence to be supported by ZOG. However, a definite synergy is possible between them. The SORM can be viewed as consisting of generic plans to carry out those tasks on the ship that are recurrent or sufficiently important to rate prior analysis. Thus, whenever dynamic planning encounters a task covered by a generic plan already in the SORM, it should incorporate that into the plan. The plan may need adaption to local circumstances, but the generic plan should still be the starting point. The reflection of this in ZOG is two-fold. First, there will be the capability to find the relevant parts of the SORM and instantiate them into a plan, thus making plan building easier and more reliable. Second, the rational structure adopted for the SORM itself is entirely in terms of generic plans for accomplishing well defined tasks. Thus, the relevance of the SORM to planning is to be maximized.

3.2. The hardware/software system

ZOG on the Carl Vinson requires a base of hardware and software. However, the requirement is actually stronger, namely, that the system be an instance of the newly emerging, high-quality interactive distributed systems. This was an essential aspect, right from the initial conception by Capt. Martin. The distributed system with high bandwith local networking is dictated not only by stations being distributed throughout the ship, but by the need for consistent rapid response at all stations and by the need for redundency. Attaining a general test bed requires that the system not be a stand-alone ZOG system. It should be capable of accepting a wide range of emerging, interesting programs from the ONR research community. Finally, embedding ZOG within such a larger system goes some way toward providing the extra power and generality that is necessary to make the system work initially and to permit it to develop gracefully.

It was our requirement that ZOG be done on the same hardware system to be used initially by SPICE, which turned out to be PERQs. There were three primary, and interlocking, reasons for this. First, SPICE is an instance of the hardware/software system needed for the job. Though only one of several such systems under development around the nation, it does represent our best efforts at the right detailed design given the current constraints of the art. Second, our continued interest in doing research on ZOG requires that it must evolve into the SPICE world. Thus, using a different hardware/software base would mean creating a separate line of development that we didn't want to be engaged in. Third, the ZOG/VINSON project is exceedingly small in terms of manpower and the time schedule is exceedingly tight. Thus, integrating with the operating system level software being constructed for SPICE was absolutely essential to getting the job done. Although this software is still under development, the degree of integration and knowledge existing within the department made it a relatively (though not completely) safe proposition.

The software task for the project starts with the SPICE81 software (to be called SALT, when it separates from the path of further SPICE development), which consists of a kernel operating system, file system, network software, and basic display package. ZOG will be coded for the PERQs within this environment in Pascal, which is currently the main language on PERQ. Recoding ZOG in a standard higher level language, from its present L* form, will increase speed and maintainability, but will decrease flexibility. It is a necessary step that must ultimately be taken for use in a non-research environment. However, the task is not simply to recode a fixed version (more precisely, the next iteration). First, ZOG must be made into a distributed system, with some portion of the ZOGnet residing on each local file system. Second, the higher quality PERQ display will be exploited by displaying two ZOG frames simultaneously. Finally, there will be a multitude of agents, written in

Pascal, which will continue to increase as we adapt the system to the actual range of tasks it must perform. Thus, the task includes the creation of an operating environment for these agents so that agents can be added and modified continually throughout the life of the system.

An interesting secondary system has emerged as the project has proceeded. There is a requirement for hardcopy of material in the ZOGnet. This occurs for three independent reasons. First, existing operations rely heavily on hardcopy, and it would be foolhardy to create a system that was locked into softcopy only. Second, the number of workstations (about 30) is far too few to rely exclusively on terminal access. For instance, copies of the SORM must exist at places other than workstations. Third, the system provides a way of creating written documents; obviously, to exploit this capability requires being able to create hardcopy. Thus, a document production system is necessary, in addition to ZOG. The current such system being used in the CMU environment is *Scribe*, originally developed as a PhD thesis here, and now being marketed by Unilogic. The solution for the project has been for the Navy to contract with Unilogic (via Three Rivers Computer Corp.) to provide Scribe on the PERQ and to obtain a Xerox 5700 laser printer aboard ship as the printing device.

3.3. Status of the project

Currently, the ZOG VAX, which was working for some time before the project was started, is being used for the development of the ZOG/VINSON ZOGnet and the agents. There are several communication lines into the VAX (4 at 2400 baud and 1 at 4800) from the Precommissioning Unit at Newport News. Much of the ZOGnet is being built from there, by Navy personnel and some shipyard personnel. Two officers of the Captain's staff, Lt. Mark Frost and Lt. Hal Powell, are in residence at CMU and participate in all phases of the project (ZOGnet building, agent programming, general design and review). The ZOGnet on the VAX contains about 15,000 frames and is growing at about 1500 frames per month. Thus, the development of the higher level components of the system is not constrained by the basic hardware/software development. As to the latter, there are now adequate PERQs in the environment (24 in July 81). The development of the SPICE81 (SALT) is scheduled for early Fall. We are beginning to develop the Pascal version of the system, mostly agents but some of the central ZOG system as well.

What would a project be like without a time schedule, and one that looked impossible to attain? Figure 3-2 shows ours. It is tied to the schedule for commissioning and shakedown cruises of the USS CARL VINSON, which is proceeding on schedule. We have targeted for a PERQ network running on the ship by Feb 82, which may seem unrealistic until one understands that this is merely a minimal

operating version and not a finished, polished system. We have allocated the full year between Feb 82 and Feb 83 to work out the kinks in the system and modify it to fit the needs of the ship as borne out with actual experience. That will also be the period of transferring maintenance responsibilities away from the CMU ZOG group to another group whose resources (and temperament) are better suited to the task.

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    |
82  |
    |- Feb ---  Ship's commissioning
    |           PERQ network running on board
    |           Begin transfer of maintenance responsibility to Navy group
    :
83  |
    |- Feb ---  Ship's deployment
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 3-2: Project Timetable

4. What makes an application succeed?

We have sketched a computer science application in midflight. It is, without doubt, ambitious with the attendant high risks. Moreover, it has none of the tentativeness of the classical sequence, in which research (here, ZOG) leads to an intent to apply, which leads to some development, then to a prototype system, and finally to real application. While we think some specific features of the ZOG/Vinson project go far toward assuring success and ameliorating the risk, our more general feeling is that many computer science applications show this same direct jump from research in mid-stream to full blown application. Let us first enumerate the special features of this application and then place it against a more general background.

Robustness. One critical issue is the robustness of the system. This applies to all levels in Figure 3-1, not just to the hardware and operating system software, where issues of reliability and maintainability are well understood. For us, the key issues of robustness occur with the ZOG system itself and with the mode of operation of the management system. These are the levels that strongly reflect the research character of the project and where the danger of an overly rigid, unresponsive

and flakey system is most acute. ZOG itself is quite robust, having been in operation for several years in a slowly evolving form. Compared with other types of research interfaces, say those for natural language, it is extremely simple. As important as the state of development of the ZOG system is its basic style of operation, which can be called *semi-automatic*. The data of the system (the frames) are entirely in human readable and manipulable form. The types of manipulation are open, being governed by a general purpose ZOGnet editor. Thus, novel or modified tasks can be still be performed, though perhaps awkwardly and tediously, even if they lie outside the capabilities of existing agents. Thus, the system starts with extensive manual operation and evolves towards higher degrees of automatic action as agents are created.

Cooperative definition. The design of the total system does not rest exclusively with any one group, nor do there exist well-defined specifications to mark the interface between the Navy and the ZOG group. In fact, the total operation can be best summed up as *cooperative design*. For instance, the management structure to be used in the system is largely the creation of the Captain and his immediate staff. The design of a ZOGnet plus agents to accomodate it has evolved to a great extent in joint sessions of Navy and CMU personnel. The content of the data base, where the details of the design are made manifest, is carried out remotely from Newport News, under the supervision of the two members of the Captain's staff who are stationed at CMU. There are dangers in such cooperative design. If serious disagreements arise it will not be easy to pin down responsibilities. But much more important is the gain, namely, that the system is one that is genuinely responsive to the needs of the users. This provides a major hedge against a notorious type of failure of technological application in the military -- systems get designed, built and delivered that are totally irrelevant to actual needs.

Evolutionary design. The task to be performed by the system aboard the Carl Vinson is not well specified. Managing the ship is of course similar to managing other large Navy ships. However, the situation is novel precisely because of the innovations being attempted by the project -- both in the concept of managing and in the means through which it occurs. Further, the management functions encompassed by the system are but a small portion of the total management of the ship, and will interact with it in unknown ways. In consequence, an attempt to engage in the standard sequence of prior study, specification, and then implementation runs great risks, even though it would help assure that at one point in time (namely, when specification is complete) attention has been paid to all that is known about the proposed application. Instead, the system design is proceeding in an evolutionary way, in which a large amount of adaptation is both expected and occurring. This approach has its own risks, of course, though due to the cooperative aspect of the design, some of these are less than they might otherwise be.

Parallel development. One standard difficulty in development projects is that they are sequential. Each stage of development must be accomplished before the next can begin. Delays are cumulative so that final stages suffer serious time compression. The ZOG/VINSON project has been able to institute parallel development in three ways to ease this difficulty significantly. First, ZOG was running on the VAX from the start. Hence the development of the higher levels -- ZOG, the application structure and the content of the ZOGnet -- are proceeding in parallel with the basic hardware/software development of PERQ, SALT and ZOG on PERQ. To attain this parallelism we have paid a significant price in not making the initial ZOG/PERQ system nearly as advanced as we would have liked. Second, the Vinson has a Wang word-processing system which they are using for limited on-line P&E experimentation prior to the ZOG P&E development. Third, because the Navy is building the ZOGnets directly from Newport News, the stage of assimilation of the system by the user is proceeding over the full life of the project. Thus, there will be many Navy ZOG users by the time the ZOG/PERQ system goes to the carrier. This is especially significant, because the actual transfer of a system to the user is a major Achilles' heel of application.

Deliberate narrowness. A key conservative decision was to limit the project scope to two specific management functions. ZOG is a general purpose interface and is capable of peforming over the full range of management functions. Indeed, ZOG might apply to other management functions even better than to the two chosen. Coupled with the generality and power of the PERQ/SALT base, the temptation is almost irresistible to design the ZOG/VINSON system to be a general purpose management information system. We have resisted. Only by concentrating attention on a few aspects will the system have any appreciable chance of success. The larger potential is there to be realized if the system is successful -- we all agree on that. Indeed, the notion of a test bed implies such general use. But the ZOG/VINSON system, including the PERQ/SALT base, must be justified and evaluated on the more narrow base. This implies a trade-off: increasing the chance of final success on a narrow scope versus the risk that the total design is too restricted by the narrow initial perspective. We think the trade-off is a good one.

Organizational commitment. A final important feature here is the nature of the organizational commitment. The initiation of this project lies with Captain Martin and the Navy. The Captain is in command, so that the application, with its problems and its possibilitites, receives appropriate attention. One important consequence is that the project tends not to be oversold by the scientists involved (i.e., ourselves). Continuous overselling by committed system developers often occurs throughout a project when the user remains skeptical even though he was sold originally. We are certainly not immune from overselling, since we are committed to the success of the project. But the

balance is shifted in the right direction, toward a realistic assessment of the project and its status.

Remaining risks. This listing of features of the project that work toward its success is not intended as a claim to guaranteed success. Rather, the features delineate the positive elements that make it possible to tolerate the risk involved. The factors that stand on the other side of the ledger are only too familiar, and attend all such endeavors: an extremely small development organization; tight and unforgiving deadlines; and many unknowns about the ship, the task and the users. These last factors, of course, arise in part from the underlying theme of this paper -- moving directly from research to application.

Typical computer science applications. We have been describing the ZOG/VINSON project, going to some pains to list favorable features. Do these make the project as a whole unique? We certainly do not think so. In fact, our point is that direct application of research results is a frequent, though hardly universal, pattern for computer science. It arises, rather obviously, from the discrete, linguistic character of software, which makes programs fundamentally understandable and reproducible, so that others can get the program to do the same things the authors did. The situation is just the opposite from that of the lion trainer -- in which only he can get the big cats to perform. Certainly software can be either good or bad, reliable or buggy, complete or incomplete. And the field is developing a notion of professionally engineered software. But it remains the case that much commercial software is on the down side of all these comparisons, and some research software is on the up side. We think that situation will remain true, permitting the short-circuiting of the full application development cycle. We also believe that many of the favorable features that characterize the ZOG/VINSON project also occur frequently in other attempts at direct application.

5. What makes an application interesting?

Researchers, just because they happen to commit to a real application, do not thereby cease to be researchers, especially if they continue to pursue their work within the same research environment. They continually attempt to derive scientific substance from the application itself. This is certainly not always easy to do, nor even possible, if too narrow a notion is taken of what constitutes a scientific result. By their nature, applications are case studies, usually without adequate control groups and without the time or freedom to shape the application to scientific ends. But if a scientific result is taken to mean something that feeds into the pool of knowledge that scientists draw upon in doing their work, then there is ample scope for scientific results to emerge from application projects. The ZOG/VINSON project provides a good illustration.

In dealing with a case study, its special characteristics must always be kept in mind. These are often useful, but they also limit what can be learned. The ZOG/VINSON project is first and foremost a management science application, while a computer interface (here, ZOG) is useful in many types of applications. The population of users is special, actually being two populations, one an officer population and the other an enlisted man population. The entire application is conditioned by the underlying *PERQ-Ethernet technology*, which is in many ways liberating compared to prior technologies, but will bring its own restrictions. Finally, the entire project is being done in an operational environment with operational pressures and time scales.

5.1. Issues of scientific interest

The research issues range over all the facets of the project. We indicate a few important ones in each category.

The total system. Taking ZOG as a specific system, the CARL VINSON provides a real application against which to test its success. Because of the way the CMU computer science computational community has evolved, we were still seeking such a driving application. If ZOG/VINSON succeeds we will have that most direct of payoffs, namely, having had an impact on the real world -- indeed, on the immediate real world of our research sponsors, ONR. Beyond this, however, it also provides a principled path for evolving ZOG into the next wave of interface technology involving locally networked powerful personal computers. This evolution is necessary for us in any event.

Management science. Several features of the system could make novel contributions to management science. One is the style of management information systems, which can be described as *interface-saturated*. Unlike most systems, where the data base is an internal structure accessible through a retrieval interface, in ZOG the entire structure is at the interface, with the user and the computer manipulating and reading the same data structures. This incurs costs in terms of control and consistency; but it also provides a structure that is highly adaptive. A second novel contribution is the attempt to characterize the procedures of a large organization entirely and consistently in task-oriented terms, sufficient, for instance, to create specific plans from them automatically. This is a very high degree of rationalization. Whether it can be attained and in what way it will be used is quite unclear. The fate of overly rigid organization, like the fate of overly rigid machines, is to be evaded by the user in order to get the work done. But there is some hope here that the openness of ZOG will allow it to become shaped by the organization to become a useful tool.

The human-computer interface. ZOG, of course, is simply one exemplar of a general

philosophy of human-computer interaction. PROMIS also exemplifies this philosophy and other variations are possible. Thus, beyond the success or failure of the ZOG system as a whole, there will be a test of the philosophy of rapid menu-selection large network interfaces. In fact, the ZOG/VINSON system will provide a first test of how this philosophy extends to a multi-window system. There are other novel general features on which the CARL VINSON experience will give evidence. How self-supporting and self-teaching is the interface? Can decentralized modification work? ZOG is a clear embodiment of these general philosophies. Success or failure (if interesting) will provide some useful information about these critical aspects. In a different direction is the compatibility of network and document representations of information. Networks, with many crosslinks and indexes, are excellent structures for on-line interaction. But it is also necessary to create hard copy representations, which must fit into the usual linear structure. Being able to move back and forth would seem to be an important capability. There is much to be learned about how to do this.

The computer system. Constructing the ZOG/Vinson system requires solving several important technical problems at the hardware and software level, problems of far greater interest than the ZOG system taken in itself. ZOG on PERQs is a fully distributed data base, a topic of much current interest in computer science. It may be that ZOG permits some special approaches; these will be interesting in themselves. The use of active agents to accomplish ZOGnet functions is another novel technical problem. It is not as much in the computer science main stream as distributed data bases, though it bears kinship to what is called procedural attachment in artificial intelligence. There are systems problems in coordinating multi-window displays. Consistency enforcement (through agents) is yet another important data base issue, which attains a particular (and interesting) form in ZOG. It is part of the general problem of mixed human-computer data structures, where the computer can depend on only moderately reliable structuring of its data, because human editing intervention occurs continually. This might seem a narrow issue, but progress in the techniques for doing it might encourage a more widespread look at this mode of operation.

Human factors. We have a substantial interest in the basic human factors of user interaction with computers. In general, of course, many basic issues can be posed within any complex interface. One interest is the problem of evaluating complex systems. Standard practice is to run an experiment comparing some aspect of the system with a control version. The difficulties of this are well known--having to freeze the system, create control versions, and run many replications--though knowing these difficulties does not alleviate the costs. The solution to this problem lies in model-driven evaluation. We have been working to build up basic modeling tools [1] and have been successful in

applying these to ZOG [10] [8]. These lead to a methodology that we call *iterative-predictive evaluation (IPE)*, which attempts to evaluate by predicting the incremental changing character of a system as it develops, rather than holding it constant. The ZOG/VINSON system is ideal for working on this problem.

Systems-design methodology. There exists in all engineering disciplines a tension between the value of full planning and prespecification, and the lack of knowledge about the consequences of design. Computer science is no exception, as the controversies in programming methodology testify. The ZOG/VINSON project is an exercise in *iterative design* as opposed to thorough and detailed prespecification. It relies on constructing primitive versions of the system rapidly, with subsequent revision and growth as the experience accumulates. It depends on being able to make modifications rapidly enough to remove the features of the system that prove to be inappropriate. Such a design philosophy fits well with other aspects of the project, such as the cooperative definition and the adaptive character of ZOG itself. But it will still be quite interesting to see how it fares.

The field as a whole. It was noted earlier that the project is to produce a test bed aboard the Carl Vinson. The focus on installing the two highly specific management functions should not obscure this. The cradle within which this sits, both ZOG itself and the PERQ/SALT distributed personal computer system, is much broader than these management functions. Captain Martin already has active plans to bring a number of other research efforts to the ship, most of them requiring a computational environment of the type being provided. For instance, a group at SRI-International is exploring an artificial intelligence interactive planning system to work on the problem of spotting planes on the flight deck [16]--an extremely serious problem given the tight space constraints. If the ZOG/VINSON project helps to provide a successful test bed for work in computer science and artificial intelligence, the effort will have benefited our science substantially.

We have enumerated quite a few issues on which the experience of the project can provide information of interest to the science. Some are substantive, some are methodological, and the last one is, in some sense, institutional. The results of the project can hardly be expected to illuminate all of these, in part because such illumination will not automatically happen but will require substantial effort. Thus, the list above is more like potential; we will be quite satisfied if even a few items pay off.

6. Conclusion

In this short paper we have sketched in brief terms a rather complex application project, which has accumulated a full range of components from hardware up through the content of a large data base (see Figure 3-1). Though we have said a bit about ZOG itself, since it is the central conceptual element around which everything is built, we have not had space to discuss any of the other components: PERQ, SALT, the new software features of ZOG on PERQ, Scribe, the two application systems (SORM and P&E), and the content of the data base. Yet all these are as critical in their own way as the central features of ZOG. Instead, we have used our space to indicate some of the underlying features that give hope for the project and some of the features that make it worthwhile for researchers to engage in such an application project.

The ZOG/VINSON project is still more than half a year away from its first major hurdle: to bring up the initial ZOG/PERQ system on the ship with a primitive version of the P&E. It would be foolish to hazard more than dedicated hopes on whether we will make that hurdle on time. There will surely be several crises before we do. Yet, with all the risks and the effort demanded, it seems to us that the potential for payoffs--in all the directions enumerated above--makes it an eminently appropriate path to be following.

7. Acknowledgements

Many people have contributed to ZOG and to the ZOG/VINSON project in many different ways. We want to acknowledge their efforts. From the USS CARL VINSON: Capt. Richard Martin, Lt. Cmdr. Frank Holland, Lt. Cmdr. Mike Rowland, Lt. Hal Powell, Lt. Mark Frost, plus all the members of the crew who are helping to create and use the CARL VINSON ZOGnet. Members of the CMU ZOG/VINSON project: Sandy Esch, Charles Kriebel, Kamila Robertson, Roy Taylor and Elise Yoder. Others who have been associated with the original evolution of ZOG: Mark Fox, Marilyn Mantei, Andy Palay, Kamesh Ramakrishna, and Andy Reiner. Finally, Marvin Denicoff of ONR, who has been midwife to this project. We would also like to thank Charles Kriebel and Capt. Richard Martin for their extensive comments on the paper.

References

1. Card, S., Moran, T., and Newell, A. "The keystroke-level model for user performance time with interactive systems." *Communications of the ACM 7* (1980), 396-410.

2. Fox, M., and Palay, A. The BROWSE system: An introduction. Proceedings of the Annual Conference of the American Society of Information Science, Minneapolis, October, 1979.

3. Hibbard, P. G. Research in Personal Computing at Carnegie-Mellon University.

4. Newell, A., Fahlman, S.E., Sproull, R.F., and Wactlar, H.D. CMU Proposal for Personal Scientific Computing. COMPCON 80 Spring, 1980.

5. Newell, A., McCracken, D., and Robertson, G. L*: An Interactive, Symbolic Implementation System. Computer Science Department, Carnegie-Mellon University, October, 1977.

6. Newell, A., Simon, H., Hayes, R., and Gregg, L. Report on a Workshop in New Techniques in Cognitive Research. Computer Science Department, Carnegie-Mellon University, 1972.

7. Pople, H. The formation of composite hypotheses in dagnostic problem solving: an exercise in synthetic reasoning. Proceedings of the 1977 International Joint Conference on Artificial Intelligence, Vol. 2, 1977, pp. 1030-1037.

8. Kamesh Ramakrishna. *Schematization as an Aid to Organizing ZOG Information Nets*. Ph.D. Th., Carnegie-Mellon University, 1981.

9. Reid, B. *Scribe: A Document Specification Language and its Compiler*. Ph.D. Th., Computer Science Department, Carnegie-Mellon University, October 1980.

10. Robertson, C.K., McCracken, D., and Newell, A. Experimental evaluation of the ZOG frame editor. Proceedings of the 7th Canadian Man-Computer Communications Conference, Waterloo, Ontario, June, 1981.

11. Robertson, G., McCracken, D., and Newell, A. "The ZOG approach to man-machine communication." *International Journal of Man-Machine Studies 14* (1981), 461-488.

12. Rosen, B. PERQ: A commercially available personal scientific computer. COMPCON 80 Spring, 1980.

13. Schultz, J.R., and Davis, L. "The technology of Promis." *Proceedings of the IEEE 67*, 9 (September 1979), 1237-1244.

14. Toedt, Lt. D.C. "Shipboard information handling tools." *United States Naval Institute Proceedings 107*, 1 (January 1981), 104-106.

15. Walton, P.L., Holland, R.R., and Wolf, L.I. "Medical guidance and PROMIS." *IEEE COMPUTER* (November 1979), 19-27.

16. Wilkins, D. E., and Robinson, A. E. An interactive planning system. Tech. Rept. 245, SRI International, July, 1981.

The Maximization Process under Uncertainty*

Richard M. Cyert
Morris H. DeGroot

The theory of the firm has been developed in general under conditions of certainty and assumptions of complete knowledge. Everyone familiar with the theory must acknowledge the great contributions it has made to our understanding of the functioning of the price system in the resource allocation process. At the same time it must be recognized that firms operate under conditions of uncertainty, and that the theory of decision making under uncertainty has been well developed in the literature of statistics and economics (Zellner, 1980; Fienberg and Zellner, 1975).

There are a number of areas in which models incorporating uncertainty have been particularly fruitful. Some of these areas are oligopoly (Friedman, 1970; Shubik, 1959); statistical decision theory (DeGroot, 1970; Raiffa and Schlaifer, 1961; Savage, 1954); rational expectations (Muth 1961; Lucas and Prescott, 1971; Grossman, 1975); theory of teams (Marschak and Radner, 1972; Harris and Raviv, 1978); asset pricing models (Sharpe, 1964; Fama, 1977); job search (Stigler, 1961 and 1962; Lippman and McCall, 1976); and econometrics (Zellner, 1971). These are only a few of the subjects in which uncertainty is being interjected and the references are only a small sample from the literature as can be seen from recent surveys (Hirshleifer and Riley, 1979; Lippman and McCall, 1978). Clearly economists are trying to incorporate uncertainty and, thereby, develop models with more empirical implications than are possible with the standard models assuming certainty.

In general, the models that have been cited involving decision making under uncertainty are based on the maximization of the expectation of an explicit function of profits. Since it is recognized that firms do not specify

operational utility functions, the actual process by which firms make their decisions has remained something of a black box and has cast doubt on the actual maximization assumption itself. Conceptually, we know that there has to be some approximation to the marginal process developed in the certainty models, but the problem is to determine how this approximation is made. We propose in this paper to describe the process in general terms on the basis of observations that we have made of that process from the top management level.

This paper represents an attempt to develop a theory of the firm that incorporates the concept of decision making under uncertainty. 1/ A business firm essentially follows a sequential process in making decisions in an environment of uncertainty. It sets goals, makes decisions, gets feedback, evaluates the feedback, and makes additional decisions. It is this sequential process for the business firm that we will analyze and model in this paper.

The Firm

The firm that we postulate is one that has a decentralized organizational structure. The basic unit in each firm is a division. Each division is a profit center and is usually managed as though it were an independent firm. The firm will sell a number of different products and through its divisions will operate in a number of different market structures in which it may have different market shares (Chandler 1966; Caves 1980).

The manager of each division will usually be a vice-president, although this person's title may range from manager to president depending on the size of the firm. The division vice-president will have the power to make

pricing and output decisions as well as some capital expenditure decisions. Furthermore, the division vice-presidents, along with the chief executive officer of the firm, will constitute the management coalition of the firm, although this coalition may also contain other executives depending on the nature of the business and the characteristics of the individuals involved (Cyert and March, 1963).

The firm in our theory operates under conditions of uncertainty. Thus the chief executive officer and his managers must make decisions about events whose occurrence is uncertain. Uncertainty as we use it might be described by a stochastic model in which the probability distribution of the observable random variables depends on the unknown values of various parameters. [2/] Under these conditions, the true values of the parameters cannot usually be learned with certainty. Thus, there will alway be uncertainty facing the firm as it makes decisions because of both the stochastic nature of the variables to be observed and the unknown values of the parameters. The description of the firm that we have presented is appropriate for most corporations and, therefore, our theory is meant to be general.

Developing a Plan

The management of the firm attempts to develop a plan from which some specific operational steps will follow. In order to develop such a plan, the management must be concerned with three sets of conditions: (1) conditions inside the firm, (2) conditions in the industry and, in particular, in competitive firms, and (3) conditions in the economy as a whole.

Let X_j denote a vector of observable variables that describes the state of the firm in month j, where j ranges over the period under study. The components of this vector are variables such as net sales, net profits, cash flow, wage levels and labor conditions. These variables are considered by the firm as it develops its plan, and their actual values each month provide useful feedback to the firm.

As the plan is developed it is also necessary to consider variables that describe the state of the industry and, particularly, the state of competition. Let Y_j denote a vector of observable variables such as prices, profits, sales, plant closings, and new investments of competitors, and variables describing an industry's condition such as price indices, industry output, industry inventories, and capacity utilization.

Finally, let Z_j denote a vector of observable, external variables in month j that describes the state of the economy as a whole and that could affect the firm's plan. This vector will typically include macro variables in the political, economic, and social spheres such as the gross national product, the unemployment rate, the inflation rate, and interest rates. In addition, summary figures of housing starts, new railroad car orders and the like, that might be vital for a particular firm, could also be included.

Establishing Targets

In the planning process management selects certain key variables, which we shall call target variables, whose actual values during the year reflect the progress of the firm (Holt, Modigliani, Muth and Simon, 1969). The management selects specific target values for these variables for the firm as a whole and for each of the divisions. [3/] The plan consists of

these target values specified for the firm and each division on a monthly basis and, when appropriate, for the entire year. The target values are the goals for the firm. Since these goals are believed by management to have a reasonable probability of being attained, the plan can also be regarded as a prediction. In effect, a predictive distribution is established for the X, Y, and Z vectors for the year, which, in turn, induces a predictive distribution for the target variables. Based on this distribution, management with the approval of the board must select the specific values that will constitute the plan.

The target values established are a compromise generated by an interactive process between the board and the management. Management wants to have high prior probabilities of achieving any targets that it selects. The board of directors wants target values that show a significant rate of growth for the firm. Thus the management tries to establish lower targets and the board tries to establish higher targets. Although the target variables may vary somewhat from firm to firm, we will specify some that are generally used in order to illustrate our theory. We shall consider six targets for the firm as a whole: (1) net earnings per share, (2) net dollar sales, (3) cash flow, (4) return on investment, (5) return on stockholders' equity, and (6) new orders received.

(1) Net earnings per share (EPS) is equal to the net profit divided by the number of shares outstanding, and represents the profit target for the firm. This variable is important because management will be evaluated on the attainment or non-attainment of its target values, and executive compensation plans are frequently tied to the degree of attainment.

(2) Net dollar sales are important because of their relationship to market share. The firm in our theory does not know its demand curve with certainty and gains information about it through the trend of net sales. Net earnings in the future are a function of this trend, as is the long-run survival of the firm.

(3) Cash flow is equal to net profit plus depreciation and deferred taxes, and is crucial to the firm since the ultimate measure of the success of the firm is the amount of cash it generates. The flow of cash gives the firm information about the need for short and long-term borrowing and thus becomes another indicator of the firm's overall well-being in the fact of uncertainty.

(4) Return on investment represents the proportion of the money invested in the firm by owners and lenders that is being returned as profit. It gives information to the firm that is important as a measure of quality of performance, and in addition, its value is one criterion of whether the firm should stay in business or not.

(5) Return on stockholders' equity has many of the same characteristics as return on investment. By focusing on the return on stockholders' equity the firm's management recognizes its responsibilities to the owners.

(6) New orders received give a measure that, like net sales, enables the firm to determine how well it is doing with respect to the future. It is a significant variable because it gives the firm information about changes in its demand curve.

The monthly target values for these six variables are important because they are related to the targets for the year. If the target value

for a variable is attained each month, the firm will reach its target for the year. Thus, whenever the actual value for a variable is less than its target value in any month, the firm is concerned.

The division targets are essentially the same as those for the firm as a whole. However, the return on stockholders' equity does not make as much sense for the division as for the firm and is replaced by a target variable representing the ratio of net profit to sales. This variable gives the division a measure of the effectiveness of its pricing policy in producing profit since it shows the amount of profit per dollar of sales revenue.

Thus, in an environment of uncertainty, the firm and its divisions establish a set of target values. Once the plan is completed, the firm must choose specific strategies to attain the target values and then wait for feedback to see if changes in strategy are necessary.

Comparison of Actual and Target

This feedback is obtained from the monthly financial statements for the divisions and the firm as a whole. Let W_j denote the vector of the actual values of the target variables achieved in month j and let t_j denote the vector of corresponding target values. Each month the financial system generates the actual values W_j, and the comparison of W_j and t_j is made. If $W_j \geq t_j$ in the sense that each component of W_j is at least as large as the corresponding component of t_j, then the firm is meeting its targets. However, we will discuss below conditions in which the firm might take action if it is exceeding its EPS target by a significant amount.

If $W_j \not\geq t_j$, then some targets are not being met and the firm enters the analysis phase of the process. An explanation must be found for the

deficiency of an actual value from its target, whether the target relates to a particular division or to the firm. This explanation must be presented in terms of the present and recent past values of the vectors X, Y, and Z. The search for an explanation depends on the relationship between the particular target variables being considered and the variables composing X, Y, and Z. If no satisfactory explanation is found, the firm takes no new actions but rather waits for more information during the following month. The process of searching for an explanation is essentially a process of making inferences from a body of data X, Y, and Z, and continues until an explanation is found with a high enough probability to warrant acceptance. Once an explanation is accepted the firm enters the next phase of the process, namely, the control phase.

Control Actions

If the explanation indicates that the deficiency in the W_j values is due to random factors that are essentially transient in nature, no action will be taken. Frequently, however, even when the explanation lies in the components of Y_j and Z_j, the firm may be able to take internal actions designed to bring the values of W_j into control, that is, to make $W_j \geq t_j$ for future values of j.

Let U_j denote a vector whose components are the various control actions that might be used by the firm in month j. Clearly U_j has many components. Some actions affect the firm's interactions with the market, such as price changes, marketing policy changes, mergers and acquisitions. Others relate to the contraction of the firm's operation, such as closing plants, selling a division, or reducing the labor force. Basically, the first set is designed to increase revenue and the second to reduce cost.

In the usual control models, a cost function drives the model, but that function is often chosen to have a canonical form, more for its mathematical convenience than its relationship to reality (see, for example, DeGroot, 1970). The specification of the cost function for the control process in the theory of the firm has the same difficulties. Consider, for example, the total earnings-per-share target. Management chooses a particular target value for a variety of reasons. Some of these may be personal and relate to executive compensation plans. Others may be professional, since the achievement of the goal is a measure of the quality of the management. Still others may involve the concept of responsibility to the shareholders. It is difficult, therefore, to give meaning to the notion of the cost of falling short of the target.

On the basis of the explanation for the target not being met, the management must decide what actions to take in order to make the actual values meet the targets in future periods. Two aspects of each action that must be considered by management are the length of time required to take the action and the length of time for the effects of the action to become apparent. As examples, we will describe five commonly used control actions: (1) price changes, (2) mergers and acquisitions, (3) contraction, (4) sale of some parts of the business, and (5) changes of management.

(1) Price change is, of course, the primary action that has generally been considered in economic theory. The firm in considering this action goes through the kind of reasoning that has generally been portrayed in oligopoly theory. The reaction of competitors is of major concern. Of equal importance, however, is a judgment about the position of the demand curve based on the information flowing from net sales and new orders. The

advantage of a successful price change is that it takes effect quickly and the firm, therefore, can be brought into control relatively soon.

(2) Mergers and acquisitions are part of a longer-run set of actions designed to make $W_j \geq t_j$. This form of control action is used when the explanation indicates a structural deficiency that results in the firm's inability to attain its targets.

(3) Contraction as a control action includes such activities as closing plants, reducing the labor force and, hence, the output, eliminating certain products completely and similar actions. These actions are taken, generally, when the firm believes it cannot affect the market and must respond by internal changes. The aim of the control is to reduce expenses proportionally more than revenue.

(4) Selling parts of the firm that are losing money is a common. longer-run type of control. Again this method tends to be used when the explanation indicates that fundamental problems in the structure of the firm are preventing it from attaining control. Frequently these segments of the business have been retained for a period of time while they are losing money because the future prospects are bright. At some point the firm makes a decision to sell. That decision will be made when $W_j \not\geq t_j$ for a number of months and the explanation leaves the firm with no other action that can be taken to make $W_j \geq t_j$.

(5) Changing management tends to be a last resort. Such action, obviously, follows an explanation that leads to the inference that management is at fault. Generally this action can take place immediately and usually a replacement from within the organization is available. Thus this action can have quick effects in making $W_j \geq t_j$. The management

changes might be at any level in the firm where the unit involved could have a significant effect on the target variables of the firm.

These five control actions are only a subset of the total number of actions that might be taken, but they are the most important, we believe, and the ones most frequently taken. The objective of management is to select a value of U_j that will make $W_j \geq t_j$. It must select appropriate levels of the available actions to achieve this objective with the least expected cost over the entire planning period.

We have discussed being out of control only in terms of some of the elements of W_j being less than the corresponding elements of t_j, but there are some control actions that may be taken when $W_j \geq t_j$. In particular, the firm is sensitive to the amount by which the actual EPS exceeds the target. The primary reason for this caution is that the management wants to show steady growth. This objective is desired because steady growth of a given percentage is an indication of good management and because it is believed that the stock market places a high value on steady growth. All other things being equal, management would prefer two years of steady growth rather than one of great growth and one of relatively low growth. The firm tries to reduce profits that will push it far beyond its EPS target by putting more funds into contingency reserves of various kinds. A general contingency reserve is not allowed, but it is frequently possible to reserve for plant damage or to develop reserves for unemployment insurance or workmen's compensation. Thus, rather than allow the actual EPS to be significantly in excess of the target, the firm will increase its reserves and reduce profits in a particular period.

The Control Model

We shall now present a general control model that incorporates the concepts and the notation that we have discussed in the previous sections. Let X_0, Y_0, and Z_0 denote the values of the observable vectors X, Y, and Z in the final month of the year before the one being studied, so X_0, Y_0, and Z_0 are specified initial values for the given year. Also, for $j = 0,1,\ldots, 12$ let $V_j = (X_j, Y_j, Z_j)$ and let $X^j = (X_0, X_1,\ldots, X_j)$, with similar definitions of Y^j, Z^j, and V^j.

The random vectors $\{V_j; j=1,\ldots, 12\}$ evolve in accordance with a controlled stochastic process that we assume can be represented by a system of equations having the following form:

$$V_{j+1} = f_j (V^j, U^{j+1}, \theta, e_j) \quad j=0, 1,\ldots, 11 . \qquad (1)$$

In (1), f_j is a vector-valued function, $U^j = (U_1,\ldots, U_j)$ and the value of the control vector U_{j+1} is chosen after V_j has been observed but before V_{j+1} is observed, θ is a parameter vector whose value is fixed throughout the process but unknown, and $e_0, e_1,\ldots, e_{11}$ are unobservable random error terms having a specified joint distribution. (Any unknown parameters in the joint distribution can be regarded as components of the vector θ.) The parameter θ represents all of the economic variables that are relevant but unobservable. The observations in any month j will depend on θ and the effects of the random shock e_j. (In an example below we let θ represents possible changes in the demand function.)

The assumption that θ is a fixed vector throughout the entire process is not restrictive. If the value of the parameter vector in (1) may change from month to month, we can let θ_j denote the value in month j and let $\theta = (\theta_0, \theta_1, \ldots, \theta_{11})$. In this case, f_j would depend only on the jth component of θ. Similarly, if the value of θ changes from month to month in accordance with some stochastic process, we can regard the stochastic part of these changes as part of the random error process $(e_0, e_1, \ldots, e_{11})$ and regard any fixed unknown hyperparameters in this process as part of the vector θ.

It should be emphasized that the initial conditions of the process will typically be described by a much wider history of the process than merely the value V_0. It is assumed that this full set of initial conditions is known and fixed at the beginning of the year, and it is not included in the notation.

Some common special cases of the system (1) are worthy of mention. It is typically true that the state X_{j+1} of the firm in month $j+1$ will depend on the state Y_j of the industry and the state Z_j of the economy as a whole, as well as on the control U_{j+1} and the state X_j of the firm itself in month j. If the industry we are studying is oligopolistic, then the state of any one firm will affect conditions in the entire industry, and Y_{j+1} will also depend on all four vectors X_j, Y_j, Z_j, and U_{j+1}. On the other hand, in a competitive industry, since no single firm can affect the state of the industry as a whole, the value of Y_{j+1} will not be affected by X_j or U_{j+1}. In this case the first two components of the equation for V_{j+1} are of the form:

$$X_{j+1} = f_{1j}\,(V^j,\ U^{j+1},\ \theta,\ e_j)\ , \qquad (2)$$

$$Y_{j+1} = f_{2j}\,(Y^j,\ Z^j,\ \theta,\ e_j)\ . \qquad (3)$$

Furthermore, it will typically be the case that the evolution of the economy as a whole will not depend on changes in any particular firm and can be studied without reference to any particular industry. Thus, we may write

$$Z_{j+1} = f_{3j}\,(Z^j,\ \theta,\ e_j)\ . \qquad (4)$$

There is no loss of generality in using the same error term e_j in (2), (3), and (4), since each of the vectors X_{j+1}, Y_{j+1}, and Z_{j+1} could conceivably depend on different components of the vector e_j. More realistically, however, it is likely that random shocks that affect the economy as a whole would also affect the industry, and those that affect the industry would also affect the firm. Hence, it is often possible to define the vector e_j so that all of its components appear in (2), all but the last j_2 components appear in (3), and all but the last $j_2 + j_3$ components appear in (4), where j_2 and j_3 are positive integers.

If we let $W^{12} = (W_1, \ldots, W_{12})$ and $t^{12} = (t_1, \ldots, t_{12})$, then there will be a total cost to the firm over the entire year of the form $C\,(W^{12},\ t^{12},\ U^{12})$. In other words, the total cost will result from realizing the monthly values $W_1, \ldots, W_{12}$ of the target variables when the corresponding monthly targets were $t_1, \ldots, t_{12}$ and from using the monthly controls $U_1, \ldots, U_{12}$. The problem facing the firm is to choose the values $U_1, \ldots, U_{12}$ sequentially in order to minimize its expected total cost

$E[C(W^{12}, t^{12}, U^{12})]$. It is assumed that the firm's utility function is a linear function of the total cost C .

In some problems, it is reasonable to regard the total cost over the year as the sum of monthly costs, and we can write

$$C(W^{12}, t^{12}, U^{12}) = \sum_{j=1}^{12} C_j(W_j, t_j, U_j) . \qquad (5)$$

Furthermore, the cost C_j itself can often be regarded as the sum of a cost C_{1j} due to missing the target in month j and a cost C_{2j} of the control action in month j . Thus,

$$C(W^{12}, t^{12}, U^{12}) = \sum_{j=1}^{12} [C_{1j}(W_j, t_j) + C_{2j}(U_j)] . \qquad (6)$$

Finally, C_{1j} may depend only on the difference between the observed and the target values of W_j , so we may write

$$C_{1j}(W_j, t_j) = C_{1j}^{*}(W_j - t_j) . \qquad (7)$$

It should be emphasized that the components of U_j are the values of the different control actions that are available to the firm in period j . Some of these components may be binary variables, where the value depends simply on whether or not that particular action was taken. Others such as a price change, are more quantitative in nature. Still others, such as mergers or changes of management are actions with a relatively more sophisticated set of possible values. In any event, the direct cost C_{2j} of using some of these actions may be taken to be 0 . Moreover, it may be true that each of the costs C_{1j} is 0 for $j=1,\ldots, 11$, and only the cost $C_{1,12}$ of missing the target at the end of the year is positive. In this case,

only the target t_{12} is important. The monthly targets $t_1, \ldots, t_{11}$ are irrelevant and need not even be specified.

Since the effects of any of the actions may be uncertain, an optimal choice may be difficult. Indeed, it is this uncertainty about the effects of the different available control actions in any given month that characterizes the control problem facing the firm. In general, the effect on X_j and W_j of a particular vector U_j of controls will depend on the unknown value of the parameter θ. Another way of saying this is to say that the particular control action that is appropriate in response to certain target variables falling below their target values will depend on the causes of this discrepancy. These causes can be thought of as being part of the unknown state of nature that the firm faces or, equivalently, part of the vector θ. The more certain the firm is about the causes of not meeting all its targets in a given month, the more certain it will be about the effects of the different control actions, and the more certain it will be about which action will be optimal. The parameter θ represents the unknown aspects of the market in which the firm is operating and of the economy as a whole.

For example, if a firm finds that its earnings in a given month have fallen below their target value, it might consider raising prices. Two major factors will enter its decision. First, a control action that changes the operation of the firm should be taken only if there has been some change in the conditions of the market or economy in which the firm operates. Such a change would be represented by the values of certain components of θ. Thus, the firm must first decide whether such changes have taken place. More precisely, on the basis of the observed values of (X_j, Y_j, Z_j)

it must evaluate its probability distribution for the different types of changes that might have taken place. Second, in the light of these probabilities, it must decide whether a price increase, some other particular control action, or no change from the previous month's operations is optimal. Typically, once the firm has observed particular configurations of (X_j, Y_j, Z_j) over a period of several months, it will become relatively certain about which action will be optimal.

The firm must be relatively certain of which action is optimal, because the effect of choosing an inappropriate control can be severe. For example, the effect of increasing prices when market conditions do not warrant an increase will be to reduce earnings further.

For example, assume that the firm begins with a known normal distribution $N(\mu_0, \sigma^2)$ for its demand at a given price p_0. In any month, the demand curve could decrease in such a way that the demand at price p_0 follows a new normal distribution $N(\mu_1, \sigma^2)$, where μ_1 is also a known number $(\mu_1 < \mu_0)$. If the firm knew that this change has occurred, and had to select a control action, we assume that it would lower its price to $p_1 < p_0$. To complete the description of the demand conditions, we assume that when the demand has the distribution $N(\mu_0, \sigma^2)$ at price p_0, it will have the distribution $N(\mu_0 + x, \sigma^2)$ at price p_1, where $x > 0$ is known. Furthermore, we assume that when the demand has the distribution $N(\mu_1, \sigma^2)$ at price p_0, it will have the distribution $N(\mu_1 + x, \sigma^2)$ at price p_1. Thus, in each month there are two possible states of the world, $\theta_j = \mu_0$ or $\theta_j = \mu_1$, and two possible control actions for the firm $U_j = p_0$ or $U_j = p_1$. No action will be taken unless $W_j \neq t_j$.

When $\theta_j = \mu_i$ in a given month, the optimal price for the firm in that month is $U_j = p_i$ (i=0,1). If $\theta_j = \mu_0$, but the firm lowers its price to p_1, it will suffer a reduction in profit through a drop in revenue. If $\theta_j = \mu_1$, but the firm keeps its price at p_0, it will suffer a reduction in profit. Thus, the loss function in each month is described in Table 1, where $a_j > 0$ and $b_j > 0$ are given constants. As the notation a_j and b_j implies, we allow the losses due to a wrong decision to vary, in general, from month to month.

θ_j \ U_j	p_0	p_1
μ_0	0	a_j
μ_1	b_j	0

Table 1

For $j=1,\ldots, 12$, we assume that $P(\theta_j = \mu_1 \mid \theta_{j-1} = \mu_0) = \delta_j$ and $P(\theta_j = \mu_0 \mid \theta_{j-1} = \mu_0) = 1 - \delta_j$. Furthermore, we shall assume that if $\theta_{j_0} = \mu_1$ for some month j_0, then $\theta_j = \mu_1$ for all months $j > j_0$. Thus, the parameter vector $\theta = (\theta_1,\ldots, \theta_{12})$ is of the general form $\theta = (\mu_0,\ldots, \mu_0, \mu_1,\ldots, \mu_1)$ and has only 13 possible values. These values depend on whether θ_j changes from μ_0 to μ_1 during the year and on the month j in which the change occurs.

In each month j, the firm observes the demand D_j for which the distribution is either $N(\theta_j, \sigma^2)$ or $N(\theta_j + x, \sigma^2)$, depending on whether the price is p_0 or p_1. It is assumed that $D_1, \ldots, D_{12}$ are conditionally independent given θ.

Suppose that at the beginning of the process $P(\theta_1 = \mu_1) = \pi_1$. In general, at the beginning of any month j, the firm will have a prior probability $\pi_j = P(\theta_j = \mu_1 | D_1, \ldots, D_{j-1})$. On the basis of this probability, the firm must choose either $U_j = p_0$ or $U_j = p_1$, and will then observe D_j. The posterior probability $\pi_j' = P(\theta_j = \mu_1 | D_1, \ldots, D_j)$ is given by

$$\pi_j' = \frac{\pi_j \varphi\left(\frac{D_j - m_j - \mu_1}{\sigma}\right)}{\pi_j \varphi\left(\frac{D_j - m_j - \mu_1}{\sigma}\right) + (1-\pi_j)\varphi\left(\frac{D_j - m_j - \mu_0}{\sigma}\right)}, \tag{8}$$

where φ *is the pdf of the standard normal* distribution and $m_j = 0$ or x according as $U_j = p_0$ or p_1. In turn, the prior probability for month $j+1$ is

$$\pi_{j+1} = \pi_j' + \delta_{j+1}(1 - \pi_j') . \tag{9}$$

Since the firm can change its price from p_i to p_{1-i} in any period, and since its choice of price affects neither θ nor the information that it receives about θ, it is not necessary for the firm to use backward induction to determine an optimal sequence of prices. The optimal price in any month will simply be the price that minimizes the firm's expected loss for that month. It follows from Table 1 that the optimal price in month j is p_0 if

$$b_j \pi_j < a_j(1 - \pi_j) \tag{10}$$

and p_1 otherwise.

In accordance with our control theory approach, the firm will consider a price change only when a target value is not being met. If the relation (10) indicates to the firm that its price is not at the optimal value, then price will be changed. On the other hand, if (10) indicates that price is at the optimal value, then the firm concludes that its failure to meet the target is a temporary phenomenon due to random fluctuations and no control action will be taken.

Summary

Our aim in this paper has been to fill in the "black box" of profit maximization under uncertainty. We have attempted to demonstrate how the firm actually behaves in its attempts to maximize profits when the environment becomes more like the real world than is true under the usual certainty conditions.

We have stressed that the process emphasizes control mechanisms and consists of the following phases:

1. Expectations. In this phase the firm examines the economy of the areas in which it operates as well as the industries with which it does business. The end result is a set of expectations about the coming year.

2. Establishing targets. On the basis of the expectations the firm sets target values for the set of variables it deems critical for profit maximization.

3. Comparison of actual and target.	On a monthly basis the firm makes comparisons between its actual values and its target values to determine whether its plan is proceeding properly.
4. Analysis.	When the actual value is less than the target value, the firm makes an analysis to develop an explanation. This phase is essentially a hypothesis testing exercise.
5. Decision to take action.	On the basis of the hypothesis accepted the firm decides that a control action is necessary or that the situation is due completely to variables beyond its control.
6. If a control action is necessary the firm selects the optimum control and implements it.	
7. The process continues in this fashion until new targets are selected.	

Our approach has been to look inside the firm to gain greater insight into the decision-making process that it follows as it attempts to maximize profit. The essence of the process is the setting of targets as a way of developing a road map in an environment of uncertainty. Once the targets are set the firm follows a control program that is similar to the behavior described by control theory.

Thus, we see the firm as a target-setting control unit operating in a sequential decision-making mode. We would argue that these characteristics are similar for any individual or non-profit organization operating under uncertainty. There is also a strong similarity to Simon's (1955) satisficing model. Both the firm in our model and the individual in Simon's model form targets or goals and aim at them. Simon implicitly assumes uncertainty and describes a decision process that ends with the achievement of the target. In our model the firm continues to operate and will exceed the target, except as indicated earlier, as circumstances dictate. However, it will not use price increases aggressively while targets are being achieved. Aggressive actions of all kinds come when the firm is failing to meet targets.

This model, we believe, gives an understanding of the operational procedures that firms follow in attempting to maximize profits under uncertainty. Its ultimate usefulness, however, will come from incorporating the model into market analysis. The perfectly competitive model cannot include this type of firm because in that model all decisions are made for the firm by the market. In oligopoly models, however, this firm has relevance and the next step in this research should be in the direction of studying the interactions of these firms in an oligopolistic setting.

Footnotes

* This research was supported in part by the National Science Foundation under grant SES-7906386. We gratefully acknowledge the comments of W.W. Cooper, J.W. Friedman, M.E. Gurtin, J. Nadler, W.H. Starbuck, O.E. Williamson, and A. Zellner.

1. This approach utilizes empirical observations of decision-making behavior in firms by R.M. Cyert in his role as a director of a number of corporations.

2. In the Bayesian approach followed in this paper both the uncertainty about the values of the observable random variables and the uncertainty about the unknown values of the parameters can be represented by a probability distribution (see DeGroot, 1970).

3. The approach of setting targets is related to the way in which the firm operates in the Soviet economy. There, however, the targets are given to the firm rather than developed by the firm's managers. The target variables are generally different but the procedure for dealing with uncertainty is similar. See Gannick (1967) and Bornstein (1978).

References

M. Bornstein. "Unemployment in Capitalist Regulated Market Economies and Socialist Centrally Planned Economies", Am. Econ. Rev., May, 1978, 68, p. 39.

R.E. Caves. "Corporate Strategy and Structure," J. Econ. Lit., Mar. 1980, 18, 64-92.

A.D. Chandler, Jr. Strategy and Structure (Garden City, N.J.: Doubleday and Co., 1966).

R.M. Cyert and J.G. March. A Behavioral Theory of the Firm, (Englewood Cliffs, N.J.: Prentice-Hall, 1963).

R.M. Cyert, M.H. DeGroot, and C.A. Holt. "Capital Allocation within a Firm," Behavioral Sci., Sept. 1979, 24, 287-295.

M.H. DeGroot. Optimal Statistical Decisions. (New York: McGraw-Hill, 1970).

E. Fama. "Risk-adjusted Discount Rates and Capital Budgeting under Uncertainty," J. Financial Econ., Aug., 1977, 5, 3-24.

S.E. Fienberg and A. Zellner, (editors). Studies in Bayesian Econometrics and Statistics (Amsterdam:North-Holland Publishing Co., 1975).

J.W. Friedman. Oligopoly and the Theory of Games (Amsterdam: North-Holland Publishing Co., 1977).

D. Gannick. Soviet Metal-Fabricating and Economic Development (Madison, Wisconsin: University of Wisconsin Press, 1967) pp. 211-217.

S.J. Grossman. "Rational Expectation and the Econometric Modeling of Markets Subject to Uncertainty: A Bayesian Approach", J. of Econometrics, 1975, 3, 255-272.

M. Harris and A. Raviv. "Some Results on Incentive Contracts with Afflication to Education and Employment, Health Insurance, and Law Enforcement", Am. Econ. Review, Mar. 1978, 68, 20-30.

J. Hirshleifer and J.G. Riley. "The Analytics of Uncertainty and Information-An Expository Survey", J. of Econ. Lit., Dec. 1979, 17, 1375-1417.

C. Holt, F. Modigliani, J. Muth and H. Simon. Planning Production, Inventories, and Work Force,(Englewood Cliffs, N.J.: Prentice-Hall, 1960.

S.A. Lippman and J.J.McCall "The Economics of Job Search: A Survey", Econ. Inq., June, 1976, 14, 155-189.

S.A. Lippman and J.J. McCall. "The Economics of Uncertainty: Selected Topics and Probabilistic Methods." Working Paper No. 281, Western Management Science Institute, University of California, Los Angeles (to appear in Handbook of Mathematical Economics ed. by K.J. Arrow and M.D. Intriligator).

R. Lucas and E. Prescott. "Investment under Uncertainty", Econometrica, Sept., 1971, 39, 659-681.

J. Marschak and R. Radner. The Economic Theory of Teams (New Haven: Yale University Press, 1972).

J.F. Muth. "Rational Expectation and the Theory of Price Movements", Econometrica, July, 1961, 29, 315-335.

L.J. Savage. The Foundation of Statistics (New York: John Wiley and Sons, 1954).

W. Sharpe. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," J. Finance, Sept., 1964, 19, 425-442.

M. Shubik. Strategy and Market Structure (New York: John Wiley and Sons, 1959).

H.A. Simon. "A Behavioral Model of Rational Choice," Quart. J. Econ., Feb. 1955, 69, 99-118.

H.A. Simon. "Rational Decision Making in Business Organizations," Amer. Econ. Rev., Sept. 1979, 69, 493-513.

G.J. Stigler. "The Economics of Information," J. Pol. Econ., June 1961, 69, 213-225.

G.J. Stigler. "Information in the Labor Market," J. Pol. Econ., Oct., 1962, 70, 94-105.

G.J. Stigler and G.S. Becker "De Gustibus Non Est Disputandum," Amer. Econ. Rev., Mar., 1977, 67, 76-90.

A. Zellner, (editor). Bayesian Analysis in Econometrics and Statistics, (Amsterdam: North-Holland Publishing Co. 1980).

A. Zellner. An Introduction to Bayesian Inference in Econometrics (New York: John Wiley and Sons, Inc., 1971).

OUTLINE

MANAGING INFORMATION TECHNOLOGY:

AUTHORITY, ACCOUNTABILITY AND IMPACTS

Peter G.W. Keen
Sloan School of Management
Massachusetts Institute of Technology

October 1981

Current shifts in information technology and applications are straining the management process in large organizations. Data Processing has generally been technocentric, monopolistic and tactical and its impact confined to specific functions and levels.

Most companies are moving towards an "ideal" capability based on telecommunications in a centrally accessible data store and workstations which create, access and distribute information. More and more of the organization's activities are becoming computer-mediated and terminals are moving the computer into new cultures.

In service industries, the ideal capability has become a key to innovation and competitive advantage. Office Technology is perceived as the main source of long-term productivity gains. Multinational companies exploit telecommunications for coordination and control.

The impacts of redistribution of data and breaking open of information monopolies makes control of information technology an overt political issue.

Communications technology is shifting the boundary between the organization and its environment. Its impacts are interorganizational, intraorganizational, interpersonal and intrapersonal. The diffusion of Office Technology - electronic mail, teleconferencing and information access - is accelerating these impacts.

The large backlogs of development projects and growing demand for applications, together with improvements in quality and

cost of services, have stimulated the rapid growth of a free market that challenges the DP monopoly.

The existing management process is not well-equipped to handle this. Organizations largely lack explicit policies for managing information technology. Authority is ambiguous and authority and accountability frequently divided. The traditional mandates, planning process and mechanisms for coorination and participation cannot easily adjust to the increased scope, salience, and impacts of communications-based applications. In particular, the issue of central direction versus local autonomy is causing severe contention, with, in many cases, DP being a barrier to innovation.

This paper presents conclusions from two completed and two ongoing studies:

(1) Policy Issues in Managing Information Technology: (international bank)

(2) Implementing Decision Support Systems in the Public Sector (6 state government agencies)

(3) Strategies for Office Technology (multicompany cases and surveys)

(4) Human Resource Policy Issues in the Information Systems field (multicompany surveys and interviews)

The paper focusses on the issue of authority and on telecommunications and business policy. It seems clear that major changes in organizational arrangements are emerging to resolve the strains on the management process, new roles and career trajectories being created and that communications technology is already eliminating many physical and temporal constraints on organizational design.

Organizational theory has had limited visibility in research and practice in the MIS field (Management Information Systems). Most of the research emphasizes individual decision making and cognitive processes. The operating model of the organization used in MIS is one of tight coupling and of information as a coupling device. Work on implementation increasingly draws on political science as its reference discipline and introduces a pluralistic perspective. The paper discusses research direction for MIS.

1. OVERVIEW

Strains in the management process being created by shifts in computer technology and applications.

Authority:

- central development versus local autonomy
- responsibility for long-term integration
- strategy for Office Technology
- direction of data management, telecommunications and Common Systems
- pressure for explicit policies

Planning and Coordination:

- emergence of key liaison and service roles
- increase in number and seniority of actors
- growing career ambiguity in Data Processing
- shift to "end-user" development

Organizational Impacts:

- policies of data
- "computer-mediated" work
- "ideal" capability
- telecommunications and business policy

2. FOUR STUDIES:

(1) Policy Issues in managing Information Technology (case studies in international bank)

- coordination of move towards "Electronic Banking"
- development of major Common System worldwide
- impact of telecommunications on marketplace
- top management policy issues
- technology delivery system

(2) Implementing Decision Support Systems in the Public Sector (two-year study of the development, implementation and evaluation of DSS for policy analysis in state government agencies)

- information monopolies
- data processing monopoly versus free market
- information "modes" and organizational design
- impact of budget process on innovation and delivery

(3) Strategies for Office Technology

Multicompany survey of efforts to implement Office Technology across the organization and to apply information technology to the "productivity problem" (ongoing).

- strategy versus tactics
- authority and mandate
- top-down versus bottom-up:
  communication-based versus word-processing
- demand side of computing
- budgeting and pricing

(4) Human Resource Policy Issues in the Information Systems (IS) field

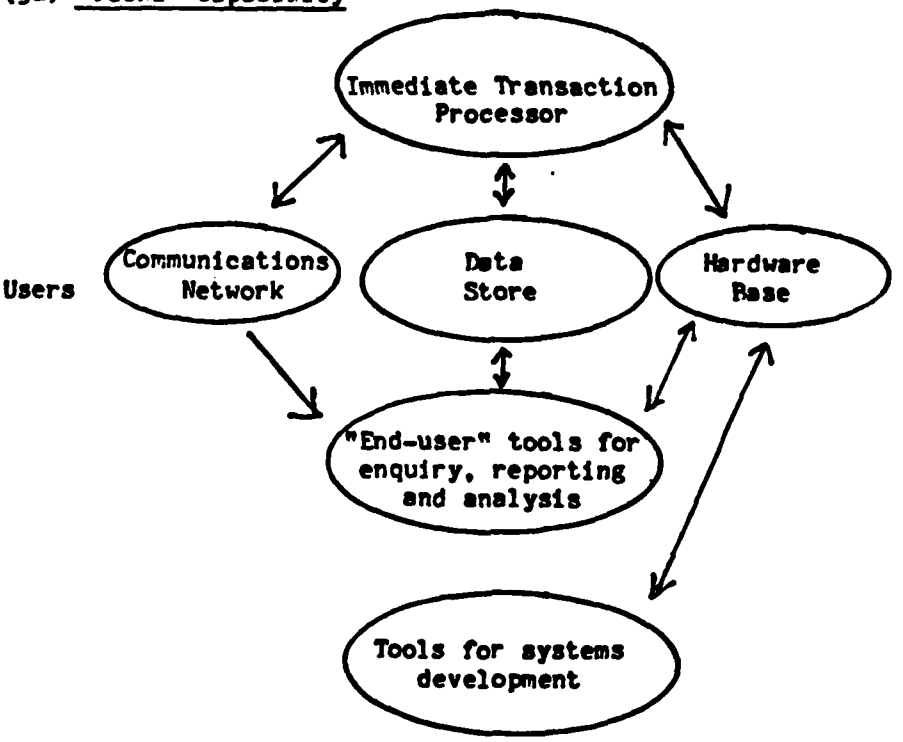
Multicompany surveys and interviews of roles, career paths and development needs in IS.

- new roles emerging: user liaison staff, technology managers, end-users support
- career ambiguity
- culture gap

3. SHIFTS IN TECHNOLOGY ENVIRONMENT AND APPLICATIONS

- "Ideal" Capability: long-term move towards integration of information resource
- End-user development: APL, functional support, IBM Infocenter, "nonproceedural" languages

(3a) "Ideal" Capability



1) Immediate Transaction Processor: Custom-Tailored, large-scale "on-line" software system. Equivalent to perfect clerk; when transaction initiated, processing completely done and "books" always up to date. In manufacturing company order entry system, in bank funds transfer system.

2) Data Store: Logically centralized library, accessed via dictionary and indices. Eliminates redundancy and inconsistency. Makes available information accessible. Standardization, clean up of existing files and coordination of data is major organizational venture. Key issues are where to begin and who is in charge?

3) End-user tools: Special-purpose packages and languages for non-programmer. Eliminate reliance on technical staff and facilitate ad hoc applications. Growing free market supply. Inefficient in terms of machines use and cost. DP concern with controlling decentralized rapidly growing use.

4) Tools for systems development. Standard programming languages (COBOL, FORTRAN) being complemented by fast development languages, application generators, "structured" design methods. Significant software bottleneck currently exists: backlogs of development projects 2-5 years, 70% of staff on maintenance of existing systems, growing end-user demand.

5) Communications network: In effect, phone lines for data as well as voices. Pushes access to the computer resource into all units of the organization (Office Technology). Links organization and its environment. In service industries, communications increasingly key to competitive advantage (banking, information services). Leads to almost every organizational function becoming computer-mediated. Complex "protocols", huge capital investment.

6) Hardware base: "Distributed" system of main frames and minicomputers. "Host" machine often manages corporate data store. Continued and dramatic price-performance improvements have transformed this from the main strategic component to a tactical one.

(3b) Free Market

DP generally internal monopoly:

- opposes or prevents use of outside services
- sets development priorities (explicit or de facto)
- obsolescent tools: retraining of technicians needed
- backlogs and lack of prior access to a free market obscures latent demand

Free market growing

- software: packages, end-user capabilities
- hardware: micros, minis, time-sharing
- communications: public networks
- data: network information services
- huge latent demand, especially for on-line, interactive systems (example: when free market for APL permitted, first-year expenditure by users $20,000,000)

- vendor strategies: bypass DP (IBM dilemma)
- micro as fifth column

(3c) Management and Organizational Issues

Long-term integration versus short-term opportunism.

- lead time
- compatibility
- certification and control: micros, local data bases

End-user demand

- free market supply
- Data Processing monopoly
- DP backlogs

Existing organizational arrangements

- hardware focus: increasingly inapplicable
- capital investment discipline: value-based, cost-based projects, based-building, portfolio criteria
- project authorization; priority setting, backlogs, Common Systems, end-user tools
- active user role: diffusion of terminals, shift in DP from manufacturing to complete business; culture gap
- allocated costs; resource-accounting, not incentive-pricing; lack of internal market, mainly ad hoc
- budget process focusses on direct technical costs: overlooks liaison, education, maintenance, R&D
- no general management tradition

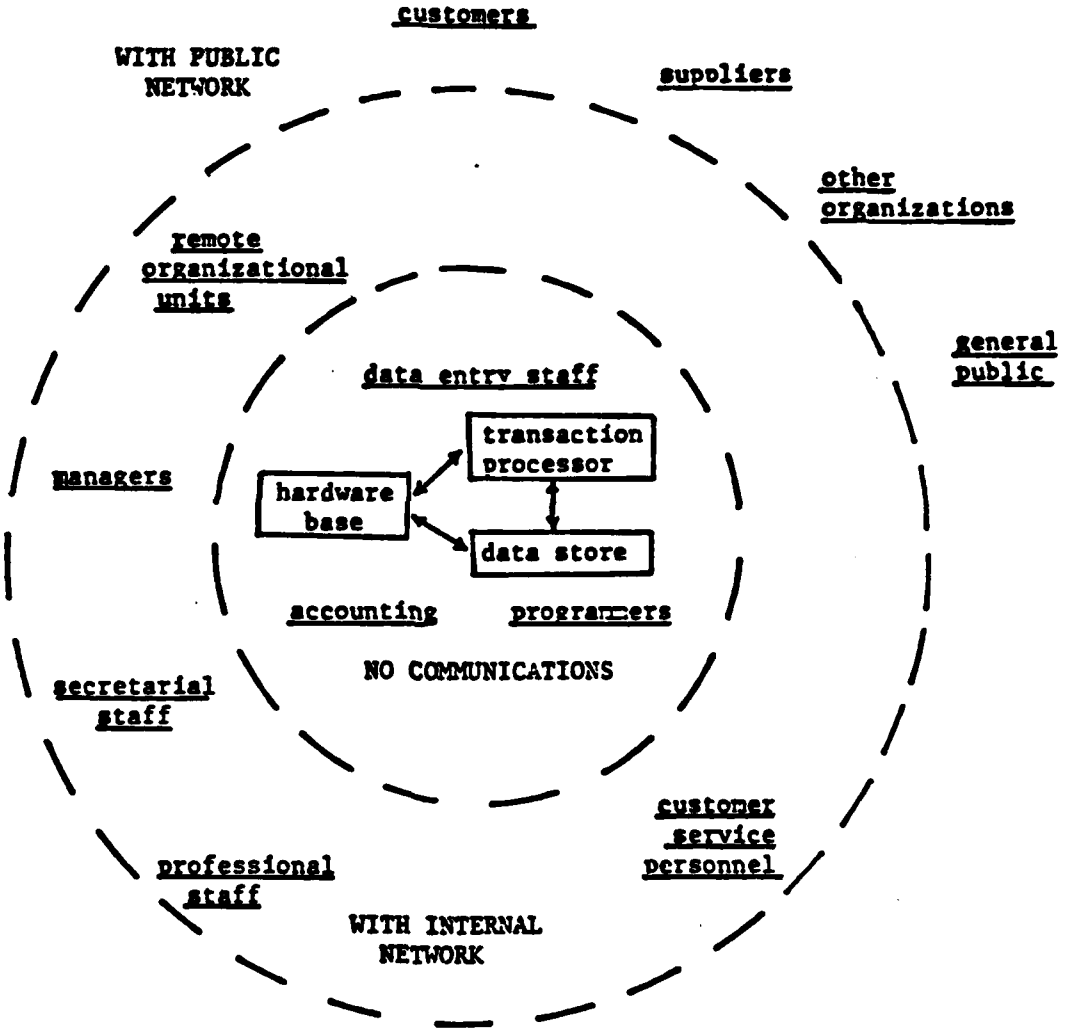
(4) Telecommunications and Business Policy

- communications and disaggregated data key to competitive advantage in service industries
- convergence and conflict of financial and information service companies

- communications and organizational design
  - role of corporate staff as message switchers (in, eg., budget process)
  - Executive Support Systems: old war-room concept; terminal access to key data; part of recentralization trend
  - decentralization-with-centralization
  - communications as substitute for travel
  - shifting boundary of organization and environment
- computer-mediated work (Zuboff)
  - text created via terminal
  - electronic mail
  - enquiry
  - abstraction of work

(4b)

Communications and the "User"



(4c) Impact of Communications on Work

Abstraction (Zuboff)

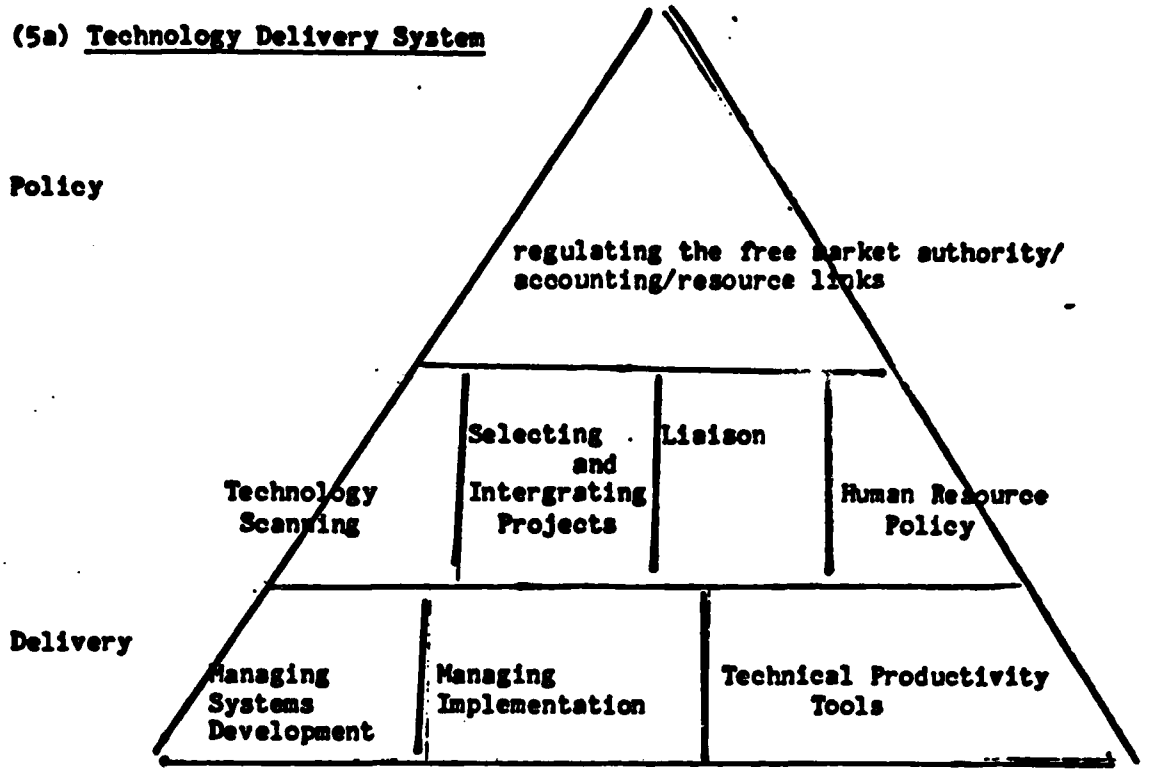
- how do you have to act?
- how do you have to think?
- are there particular forms of stress?
- what patterns of social interaction are supposed?
- how do people make sense of their new work?
- how do they think about their own productivity and value of what they do?
- how do they decipher management's intentions?

Social Interaction

- saying hello (Johansen, Vallee and Spangler)
- electronic mail and short tempers (Earnest)
- supervision (Roethlisberger revisited)
- inspection/evaluation (Dornbusch and Scott)

5. POLICY ISSUES

(5a) Technology Delivery System



- implicit policy; top management's role reactive
- no portfolio: aggregation of projects
- pressure for local autonomy
- need for central direction for communications and data management
- mandate for R&D, planned variances, Common Systems, veto
- creation of user liaison/service roles
- major shift in mechanisms for negotiation, capital investment planning and budgeting

(5b) Policy Issues

- regulating the free market
- linking authority/accountability/resources
- managing systems development

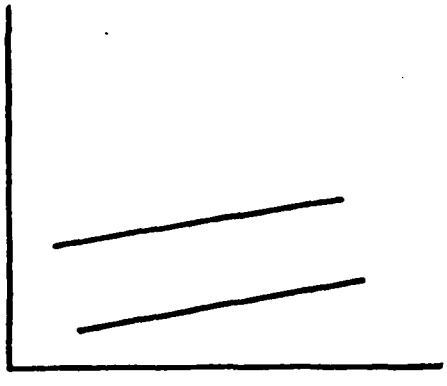
- managing organizational change
- project portfolio
- capacity planning
- network/data strategy
- intrusion of new applications: Office Technology, Decision Support Systems
- "referees" and direction from top
- committees
- reporting relationship and level of IS
- formality of disciplines

6. POLITICS OF DATA

- information monopolies; corporate staff, state agencies (Pogrow, Pettigrew, Barift and Galbraith, Keen)
- preempting policy debate
- creating information modes; problem of institutionalization
- counterimplementation

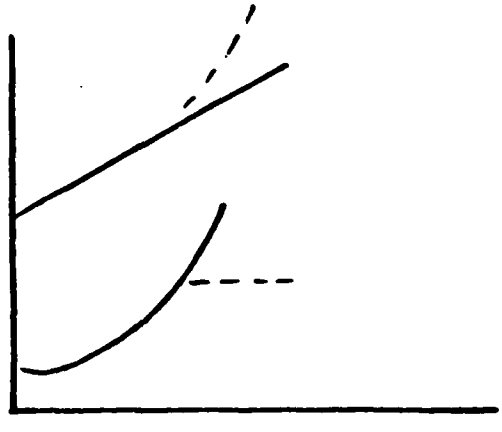
7. RESEARCH ISSUES IN MIS

- demand side of computing
  - elasticity
  - value analysis
  - premium features v. commodity
  - support needs
  - "control" marketing
- telecommunications and organizational design
  - information as coupling device
  - terminals and culture clashes
  - electronic meetings
  - "logical" versus physical arrangements
  - impact on supervision, climate, organizational identity
  - interaction organization and environment (e.g. customer service)
- telecommunications and business policy
  - ideal capability
  - service industries
  - "productivity"



manufacturing

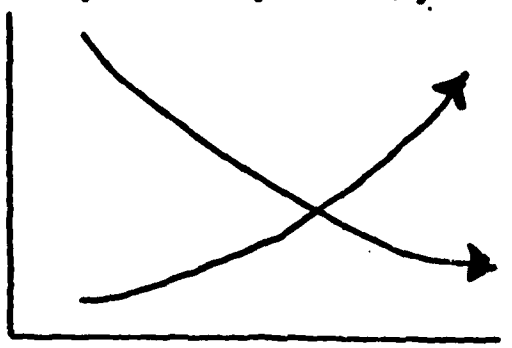
economies of scale



service

diseconomies of coordination and administration

- perceived equations of productivity



- hidden costs of decentralized computing (King and Kraemer)

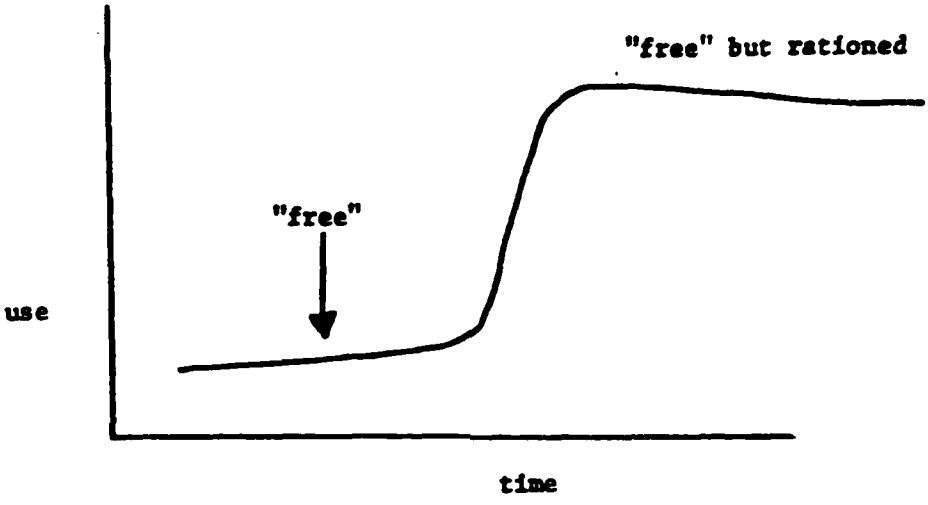
- new roles (Keen, Bronsema and Zuboff)
  - technology managers
  - liaison roles
  - career ambiguity
  - career trajectories
  - evaluation of implementers who are not builders

- leadership
  - policy impact
  - "new breed" of DP managers
  - who succeeds DP?

- economics of information (Treacy)
  - dependent variable
  - incentives
  - accounting

8. ACCOUNTING FOR INFORMATION

- purpose of accounting
- naive average cost tradition
- infrastructure mainly, fixed cost: capacity and incentive pricing
- innovation and "control"



- measuring "information"
- monitoring costs
- externalities
- objectivity versus reliability
- standard costing
- Information Economics
  - costs of processing
  - characteristics of information processor
  - value and cost not independent
  - ignores technology: cost and value depend on physical organization of files and capabilities of software that accesses the data items.

FORMULATING AND JUSTIFYING BUDGET PROBLEMS; BAD NEWS AND NOT SO GOOD NEWS*

Patrick D. Larkey and Richard A. Smith
Carnegie-Mellon University
October 1981

*Paper prepared for presentation at the First Annual Symposium on Information Processing in Organizations, Carnegie-Mellon University. We are indebted to Carol Coyne, Jeri Negler and Felicia Pratto for assistance on data collection. We are especially indebted to David Lehman for his herculean efforts on data and programming.

I. Introduction

> The budget is the life-blood of the government, the financial reflection of what the government does or intends to do.
>
> --Aaron Wildavsky
> Prior to writing his first book on budgeting

> I'll never understand why people believe this budget business is so important.
>
> --Tip O'Neill
> Speaker, House of Representatives. Prior to the introduction of the Reagan Administration budget

I.1. Motivation

Untrustworthy verbal reports are commonplace. We constantly encounter reports which we know immediately from the context or the source that we should not take at face value. Consider the following examples:

-- A politician's campaign promises or his protestations that he is not running for that office.

-- "Absolutely Final Offers" in many bargaining situations.

-- An academic's forecast of when a book or paper will be done.

-- A used car salesman's report who previously owned the car and what the "real" mileage is.

-- A defense contractor's estimates of what a weapons system will cost, what its performance capabilities will be, or when it will be ready for deployment (insert "garage" for "defense contractor" and "car" for "weapons system").

-- A husband's account of how the lipstick came to be on his collar.

-- "News" items in TASS.

-- A corporate or municipal bond prospectus.

-- Barroom accounts of athletic and/or sexual conquests.

-- Marriage vows.

-- Letters of recommendation.

-- The Marketing Manager's estimate of next year's sales.

-- A stranger's report on his golf handicap that accompanies his suggestion of friendly bet.

-- Your teenage son's explanation of almost anything.

-- Voluntary accounts of annual income.

-- A child's explanation of how the expensive vase came to be broken or the cookies came to be eaten.

-- A football coach's appraisal of next week's opponent.

The list is practically endless. You can construct your own list which may or may not agree completely with this one. But there would be a remarkable consensus on many of the entries. As Feldman and March (1981, p.175) observe, ".... Much of the information used in organizational life is subject to strategic misrepresentation."

This paper is concerned with another class of untrustworthy verbal reports: statements by chief executives or financial officers in government as to what their budget problem is in any given year. We examine how budget problems are formulated, why they are formulated as they are, how the problem formulation affects solution behavior, how budget problems and proposed solutions are reported and why the problems and proposed solutions are reported as they are. We go one step beyond idle theoretical speculation about these topics and examine, in some detail, budget formulation and justification in the City of Pittsburgh from 1940 to 1976.

We still lack satisfactory positive theories of budgetary processes, and such theories would enable us to do some "organizational engineering" on organizations and decision systems which obviously need it. Remember New York, Cleveland and Detroit? The theories are deficient on many counts. The aspects

of budgeting we examine here are among the least studied aspects of budgeting. We believe that a better understanding of these aspects is important to improving the budgetary theories. It is well known, at least in the Cognitive Psychology and Artificial Intelligence communities, that the way in which you represent a problem (i.e., the "problem space" you construct) is a critical determinant of what solutions you find and the efficiency with which you find them. Also, budgeting is a social process. When those with primary budgeting responsibility communicate their problems and proposed solutions, they simultaneously give a rationale (justification) for what they have done and what they propose to do. This rationale should partially determine the acceptability of the formulation and proposed solutions. And the knowledge of the need to provide such a rationale should be an important constraint on problem formulation and solution.

I.2. Organization of Paper

The balance of this paper consists of three parts. In the next section, we discuss budgeting as a "problem", review the literature, and describe our hypotheses. In Section III, The Pittsburgh Case, we describe our sample and data, develop a simulation of budget problem formulation, and present some descriptive data that bear on our hypotheses. In the final section, we summarize what we learned and discuss some normative implications of that knowledge. We also look briefly at the task of constructing a full, operative simulation of municipal budgetary process.

II. Budgeting As A "Problem"

> In March, 1842, the Prime Minister introduced his first Budget - an epoch-making Budget - in a weighty and lucid speech.
>
> Nothing was concealed, nothing kept back; the full financial position, in all its ugly nakedness, was laid bare. The closing year would end in another deficit of some two millions and a quarter, following on a series of deficits, amounting for the previous four years to over five millions. The coming year too, on the basis of the revenue of the previous year, would end in a further deficit of two and a half millions for the "general service of the year." In addition, the Chinese war would cost some hundreds of thousands; and there was a probability that England would be called upon to give pecuniary aid to India on account of the renewal of the Afghan war.
>
> The deficiency was no casual one; deficits had become chronic. To reduce the expenditure was out of the question. An increase of revenue could hardly be expected. Convinced though he was, that a wise reduction of duties would eventually produce a larger revenue, Peel would not, as his predecessors had done, calculate on an immediate recovery without providing for a possible deficiency.
>
> - Sidney Buxton [1894]

> Whenever there is a gap between where you are now and where you want to be, and you don't know how to find a way across that gap, you have a problem ... The process of finding a solution has two major parts: 1. Representing the gap -- that is, understanding the nature of the problem, and 2. Searching for a means to cross it.
>
> - (Hayes, 1981, i)

Prime Minister Peel's budget situation is easily seen as a "problem" in Professor Hayes' terms. Peel's gap is a projected deficit between revenues and expenditures and from this excerpt, it is not clear that Peel knows how to close

the gap. Indeed, as Peel's problem is presented with no discretion on expenditures and little discretion on revenues, the problem may have no solution.

The purpose of this section is to examine budgeting as a problem-solving process with three main aspects: formulation, solution, and justification.* Formulation is that part of the process in which the problem-solvers (budgeting officials) define their initial state (A) -- where they are now -- and their terminal state (B) -- where they want to be. The formulation (A - B) defines a solution space that consists of a set of operators (⟹) which can be applied to reduce or eliminate the gap between A and B. Solution is finding a set of operators (⟹) which solves the problem. Justification is a part of the process in which the problem-solvers (e.g., the mayor and financial officers) attempt to "sell" their formulation and solution to others (e.g., city council and the public) who must modify or approve the formulation and solution as presented and justified.

The separation of these aspects is somewhat artificial. Reformulation is one form of solution (e.g., if Peel had not reformulated his problem in 1842 it probably could not have been solved). Solutions often are an integral part of the formulation; if there are obvious solutions (e.g., leaving vacant positions unfilled) which make it possible to represent an "easier" problem, they are often incorporated in the initial formulation. And the justification phase can be seen as part of the solution because until a proposed formulation and solution is accepted, the problem is not solved.

Moreover, the budget problem is dynamic and continuous. There is one decision occasion annually, budget preparation time, when City officials are required to address the problem directly and to produce a "solution" in the form of estimated revenues and appropriations for the coming year which balance.

---

*The problem solving concepts used in this section are taken from Reitman (1967), Simon (1967), Feigenbaum (1969) and Newell and Simon (1972).

But the budget problems across years are interdependent. This year's budget problem is "harder" or "easier" in large part because of the ways in which the budget problem in prior years has been solved. Also, although solving the problem of passing a balanced budget at the start of the fiscal year is usually the most important period of problem-solving in terms of changes for a budget year, the budget is reformulated, resolved and rejustified throughout the year. Indeed, the auditors and officials are usually reformulating, re-solving and rejustifying the budget problems as much as six months after the budget year, in real time, has ended.

Throughout the discussion in this paper, we will adopt the perspective of the Chief Executive -- in Pittsburgh, the Mayor -- who is responsible for proposing annually a balanced budget to City Council in describing budgeting as a problem-solving process. It would be more correct, but extremely tedious, to describe the somewhat unique budget problems of each of the various participants in the process from budget analysts to citizens. But since many of these participants have little to do with formulation and justification, and our purpose is not to articulate a full-blown theory of budgeting in problem-solving terms, the simplified perspective is justified. An alternative would be to treat the organi-zation as a monolithic problem-solver. This can be useful strategy if all that the research requires are good predictions of the organization's behavior and modelling the internal characteristics of the organization is not essential to the accuracy of the predictions (Larkey, 1979). But the purpose of the research reported here is to understand budget problem formulation and justification.

With the above qualifications in mind, this paper examines budgeting as a problem-solving process. The initial state (A) is the budgeters' best estimate, prior to the start of the budget year, of what the budget would be if policies (tax rates and levels of service input) remained constant. They forecast revenues

and expenditures and calculate a gap between the two. The terminal or desired state (B) is a balanced budget that satisfies a variety of other constraints from within the organization and from the external environment. The solution space consists of adjustments to revenues and expenditures or more arcane ploys to remove the gap. It includes:

Increase Revenues

Raise rates on existing taxes
Introduce new tax
Raise water rates
Tap surplus
Increase non-tax revenues

Decrease Expenditures

Refund maturing bonds
Cut fixed percentage
Cut selectively

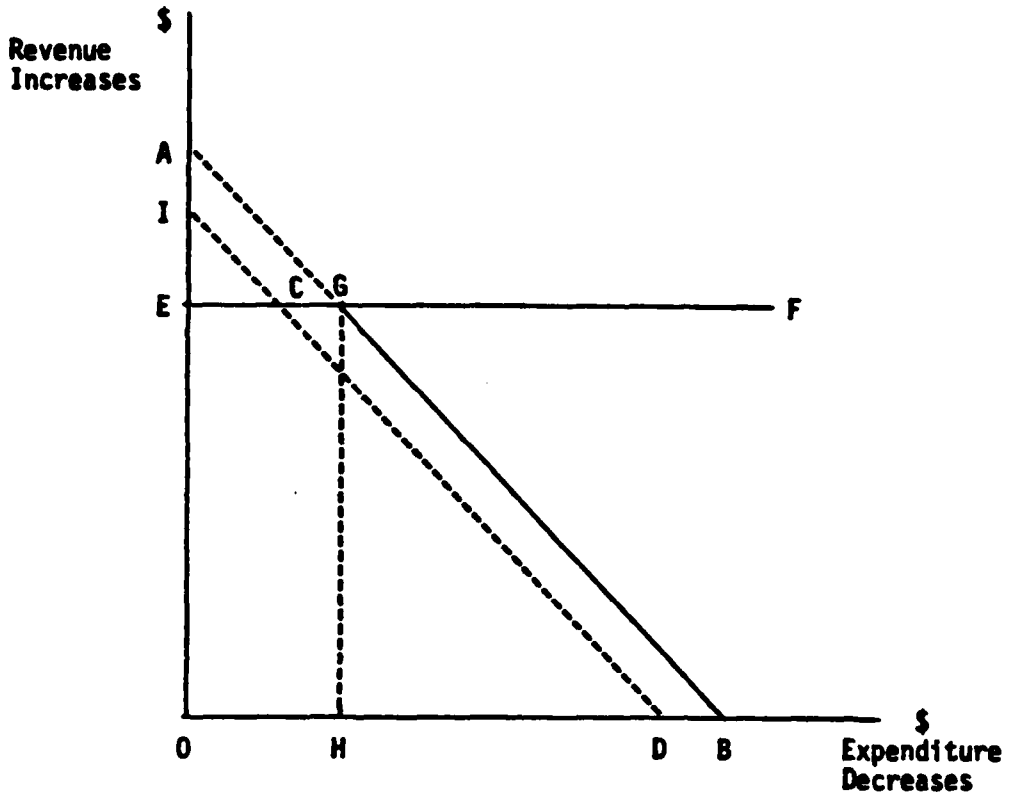
Run a Deficit

Another way of thinking about the solution space for budgeting problems is graphically. Figure 1 is one possible representation for the solution space for a problem of magnitude OA = OB. OA is the amount of additional revenue that would have to be raised to close the gap as presented with no expenditure decrease. OB is the amount of expenditure decrease that would be required to close the gap as presented with no additional revenues. The combinations of revenue increases and expenditure decreases which balance the budget are on the line segment, AB.

The solution set can be altered in a number of ways. EF might be an effective cap on the jurisdiction's revenue raising ability such as a ceiling on tax rates in state enabling legislation. The effect of the cap is to remove all points on AG from the solution set which means that an expenditure decrease of at least OH will be required to balance the budget. ID might be the "real problem" in contrast

FIGURE 1

SOLUTIONS FOR A PROJECTED DEFICIT



with AB, the "reported problem." Obviously, solving the problem on AB rather than ID leaves the government in a projected surplus position.

II. 1. Formulation and Justification in the Budgeting Literature

There is surprisingly little discussion of problem formulation and justification in the literature on budgeting. There is very little literature with a problem-solving orientation. The great bulk of the budgeting literature takes one of two approaches. First, there is the approach of the microeconomists and political scientists of the microeconomic theoretical persuasion. The earliest application of this view to Government budgeting that we can find is in the work of Pantaleoni (1883, p.19).

> We may take it that Parliament sets out to obtain the largest total sum of utility by weighing the country's wishes against its capabilities - by whatever standards both may be measured; clearly the approval or rejection of any expenditure must be based on a judgement which is the resultant of a complex of different elements or propositions. These are obtained by arranging in decreasing order the marginal utilities deriving from the various possible expenditures and then weighing the marginal utilities inherent in every combination of possible expenditures against the marginal sacrifice caused by the total taxation which each of these combinations would entail.

Another early version of this theory applied to budgeting is attributable to Pigou(1928).

> As regards the distribution, as distinct from the aggregate cost, of optional government expenditure, it is clear that, just as an individual will get more satisfaction out of his income by maintaining a certain balance between different sorts of expenditure, so also will a community through its government. The principle of balance in both cases is provided by the postulate that resources should be so distributed among different uses that the marginal return of satisfaction is the same for all of them.. Expenditure should be distributed between battleships and poor relief in such wise that the last shilling devoted to each of them yields the same real return. We have here, so far as theory goes, a test by means of which the distribution of expenditure along different lines can be settled.*

---

*Quoted in Key, (1940).

Peterson (1980) provides a modern version directed specifically at local government.

> ... A local government has an overriding set of interests that it must constantly strive to protect. Just as the nation-state wishes to maximize its power in the international system, and the private firm wishes to maximize its profits, so the local community wishes to maximize its economic well-being, its social standing and its political position relative to other local units."

Not much has changed in this line of inquiry in the last one hundred years. The mathematics behind the formulation are increasingly more elegant. The formulation continues to be irrelevant empirically and wrong as a positive theory of budgeting.

This essentially normative view of this problem, little changed in one hundred years, has the usual seeds for confusion between normative and positive that are so common to rational models. The norm is so compelling that it is hard to imagine why anyone would violate the norm. Hence, the norm becomes the only sensible descriptive theory. (See Kadane and Larkey, 1980 for a fuller discussion of the normative-positive confusion.)

In problem-solving terms, the implicit problem in the economists' formulation is an infeasible representation. The terminal state for the problem is a position of maximum utility (or whatever) and this terminal state makes the problem an Ill Structured Problem (ISP). There is no test that can be applied to the proposed solution in a finite number of steps (Simon, 1967). This strand of literature is, therefore, irrelevant to the topics of this paper.

A second strand of literature is the highly inductive literature that has regressed almost all possible transformations of budget data on all available socio-economic data. This literature essentially began with Davis et al, (1966) and has been

adequately criticized [(Padgett, 1980); (Wanat, 1974); (Natchez and Bupp, 1973); Fischer and Crecine, 1980)]. The implicit budget "problem" in this literature is to adjust appropriations program-by-program. This problem is solvable but under-specified and not very interesting absent some overall coordinating mechanism such as Crecine's "Great Identity" or a balance requirement.

One notable exception to this general inattention in the literature to budgeting as a problem that is formulated and solved is the work of J.P. Crecine. In recent work (Crecine, 1975; Fischer and Crecine, 1980), Crecine has characterized the U.S. Federal Government's budget problem as solving "The Great Identity" (Fischer and Crecine, 1980):

$$PD_t + PN_t = R_t^* - d_t$$

where

PD_t = Planned Defense Outlays in t

PN_t = Planned Non Defense Outlays in t

R_t^* = *Revenue Forecast* in t

d_t = Planned deficit in t

The explanation for why this is the federal problem is that fixing R_t^* and d_t in The Great Identity provides "Top-down" coordination for an otherwise unmanageable problem of achieving a reasonable total outlay by working on thousands of defense and nondefense programs. By first solving the identity, "these aggregate planning targets (i.e., PD_t and PN_t) may then be used as a basis for issuing "Planning Guidance" concerning specific agency budgets. They may also be used as constraints to be applied during the process of cutting requests for specific programs." (Fischer and Crecine, 1980, pp.6-7.)

Padgett (1980) offers an alternative, the Theory of Serial Judgement, to the Davis et al, (1966; 1971; 1974) incremental linear models of the federal budgeting

process which has much more of a problem-solving orientation than most of the budgetary literature. Padgett's theory of the process is a stochastic process model which he summarizes as follows.

> Given ... fixed historical referents [Previous budget estimates]... the serial judgement decision maker then decides upon a direction of search - namely, upon whether some increase is warranted both "on the merits" and "under the current fiscal climate"... Choice under the serial judgement decision strategy is simply a matter of cycling through the salient alternatives encountered along the direction of search until one of them is deemed acceptable both on the merits and under the current fiscal climate. (Padgett, 1980, pp.363-364)

Padgett is not explicit about problem formulation or about how solutions are justified. Both topics appear to be buried implicitly in two probabilities (β_i and $\bar{\beta}$) that determine whether "salient alternatives" are "acceptable" based on "merit wants" and "the current fiscal climate." $1 - \beta_i \bar{\beta}$ then becomes "a stochastic analogue to Simon's more deterministic notion of 'aspiration level,' since probability of rejection increases the higher the aspiration level."

Crecine's first work in the budgetary area was on municipal budgets and his conclusions about what the budget problems were and were not are worth quoting at length.

> It is quite clear (from interviews) that the decision-makers do not see the problem as one of optimally balancing community resources, allocating funds among functions to achieve overall community goals, and the like.
>
> The problem is generally considered by department heads as compiling a budget request that (1) assures the department of funds to carry on existing programs as part of a continuing attack or existing problems, (2) is acceptable to the mayor's office, (3) and provides a reasonable share of any overall budget increases to the department to enable it to attack new problems (if any).
>
> The mayor's problem is largely one of recommending a budget that (1) is balanced, (2) at least maintains existing service levels, (3) provides for increases in city employee wages if at all possible, and (4) avoids tax increases (especially property tax increases in the belief that increased property taxes cause business and industry to move from the city, reducing its tax base).

> ... The problem for the council is to review the mayor's budget recommendations and check for obvious errors and omissions. Because of the complexity and detail in the mayor's budget and lack of council staff, the council's options are limited largely to approving the mayor's budget. The requirement of a balanced budget means that a change in one expenditure category, for instance, implies a change in other categories and for other administrative units or a change on the revenue side of the bill - i.e., one change in the budget (by council) implies many changes which the council has neither time nor staff to consider. (Crecine, 1969, pp.38-39)

In this work, Crecine noted the close correspondence between his model of budgetary processes and a more general model of problem-solving processes.

> GPS's correspondence with the municipal resource allocation model appears to be extremely close. The goal is to transform preliminary budget recommendation totals to the revenue estimate total. A difference between anticipated expenditures and revenue is calculated and the goal of reducing this difference (i.e., eliminating the surplus or deficit) is established. A set of operators relevant to the differences of a given type is evoked - deficit elimination procedures or surplus reduction procedures, depending on whether the difference is positive or negative. The various difference-reduction routines are tested to see if the conditions for applying a particular operator are met (Is the anticipated surplus large enough to finance a minimum salary increase?). If so, the operator (raise salaries uniformly) is applied, and so on. Two conclusions emerge:
>
> 1. Our model of municipal resource allocation is consistent with the information-processing approach found in GPS.
>
> 2. A means-end analysis of the municipal resource allocation process appears to be a reasonable one, where the end is a balanced budget and the means consists of a fixed set of residual elimination routines (surplus reduction for positive residuals and deficit elimination for negative residuals). (Crecine, 1969; 210, 215-216)

It is this basic insight of Crecine's that we are using in this paper.

II. 2. Problem Formulation and Justifications: Some Hypotheses about Bias

> The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil.
>
> - Bertrand Russell

As we noted in the introduction, the purpose of this paper is to examine how budget problems are formulated, why they are formulated as they are, how the problem formulation affects solution behavior, how budget problems and proposed solutions are reported and why the problems and proposed solutions are reported as they are. Our emphasis is twofold. First, we want to examine how the stated formulation of the problem differs from the "actual" problem and, hence, how proposed solutions differ from those that are "really" necessary. Second, we want to investigate how justifications of the stated formulations and solutions increase the prospects of their acceptance by both legislators and citizens.

Our interest in strategic aspects of problem formulation and justification stems from a longstanding observation in the budgeting literature: budget officials underestimate revenues. Evidence of revenue underestimation (or "conservative bias") has been found in nations (Panteleoni, 1883), developing countries (Caiden and Wildavsky, 1974), states (Turnbull, 1967), school districts (Gerwin, 1969), municipalities (Larkey, 1979), and business firms (Cyert, March and Starbuck, 1961). The explanations for the phenomenon of underestimation include a desire to reduce the possibility of budgetary conflict (Gerwin, 1969, p.35), a strategy for hedging against uncertainty (Caiden and Wildavsky, 1974), and a desire to make subsequent versions of the budget problem in later years problems of eliminating surplus rather than closing a deficit gap (Larkey, 1979).

For similar kinds of reasons, we expect other aspects of the formulation and justifications of budget problems to be biased as well. For example, the natural counterpart of the underestimation of revenue is the overestimation of expenditure requirements. The two taken together overstate the magnitude of the real budget problem, thereby providing budget-makers with even more uncommitted resources for reducing conflict, hedging against uncertainty, or insuring ample surpluses in future years. In this paper, we investigate three hypotheses about biases in formulation and justification. The hypotheses were deduced, very casually, from prior research and our experience with budgets and budget-makers. There is no theory in this area from which these hypotheses can be rigorously deduced. They are, however, "sensible" in light of what is known about the underestimation of revenue phenomenon and, more generally, budgetary processes.

The three hypotheses are:

Hypothesis 1: Chief executives present biased formulations of budget problems to legislators and the public. The formulation consists of biased estimates of revenue and expenditure requirements and is calculated to show that the budget problem is worse than the unbiased estimates would show.

Hypothesis 2: The extent of the bias in the executive's formulation of the budget problem is conditional.

a) Budget problems are presented as worse than they "really" are when there are no strong external pressures (an impending election) to avoid certain solutions to the budget problems (tax increases).

b) Budget problems are presented as worse than they "really" are when tax changes (either increases or reductions) are proposed than when no change (especially in those years prior to a tax increase) is proposed. Chief executives, in other words, extract (in years of tax increases) and retain (in years of tax reductions) resources greater than they "really" need.

c) Budget problems are presented as worse than they "really" are when expected revenues exceed or approximate expected expenditures. Budget formulation becomes more accurate as the size of projected deficit increases.

Hypothesis 3: Chief executives present biased public justifications of the formulation of the budget problem. The bias occurs in the justifications of proposed increases in expenditures and takes two forms.

a) Chief executives single out for attention only a subset of the proposed increases in expenditures for specific programs, typically those viewed as highly desirable by the legislature and the public (e.g. public safety and parks and recreation).

b) Chief executives overstate the proportion of the total increase in expenditures that are "really" due to increases in desirable expenditures.

The a priori specification of hypotheses is both virtuous and dangerous. The virtue in the exercise is that it focuses your data collection and analytic activities making the research task more manageable. Given the wealth of possible data and analytic procedures for this research, the simplification was welcome. The danger in the a priori specification of hypotheses is that your data collection and analytic activities are focused by a particular representation of the research problem which may be wrong; you run the risk, if wrong, of not learning very much even though there may be materials of interest if you wore a different set of theoretical blinders. Given the exploratory nature of this research, we were careful to think broadly in the analysis. Indeed, one of our hypotheses (#3) was simply wrong and what we found in its place is probably our most interesting empirical result. In the next section, evidence is considered for each of the hypotheses. But prior to that consideration, we describe the data upon which the evidence is based and the method for establishing the "real" budget problem.

III. The Pittsburgh Experience

That's not an experiment you
have there, that's an experience.

- Sir R. A. Fisher*

III. 1. A Sample of One by Thirty-Seven

Our empirical work for this paper examines budget formulation, solution and justification for one city, Pittsburgh, over thirty-seven years, 1940 to 1976. The theoretical discussion above and some of the analysis was informed by prior and concurrent work on other municipalities. Although desirable, it was infeasible to do comparative work in this paper; archival data of the type we use for Pittsburgh was not immediately at hand for other cities.

The longitudinal perspective is important to the topics we are addressing. As we noted in the last section, the budget problem is perpetual; the formulation, solution and justification behavior is to a large extent continually determined as the context evolves. It makes little sense to sample single years across governments for the topics we have posed.

The thirty-seven year perspective we have in the data on Pittsburgh is sufficiently long to capture considerable variation in the external and organizational environments for budget problems. In the period studied, the city reached its peak population and saw that population decline precipitously; six mayors held office; major changes in the city's taxing power occurred as a result of state enabling legislation; and the city experienced several periods of real budgetary affluence and several periods of real budget scarcity. Whether the results that emerge from this research have any generality to other periods in

---

*Quoted in Freedman et al (1978), p.10.

Pittsburgh and to other cities and to other forms of government, is a question that can only be answered through further empirical work.

The data on Pittsburgh's budgetary process used in the analyses below were drawn from several sources. These include:

1. Actual expenditures and revenues, council approved revenues and debt and interest requirements (1940-1976) as recorded in the Annual Report of the City Controller, City of Pittsburgh, 1940-1976.

2. Mayor's estimate of revenues, expenditures and explanations of the formulation and solution of the budget problems and council approved appropriations (1947-1976) as presented in the Municipal Record...Minutes of the Proceedings of the Council of the City of Pittsburgh, Budget Address of the Mayor and the Appropriations Ordinance, Volumes 80-109.

3. Mayor's estimates of revenues and expenditures (1940-1946), and newspaper presentations of the mayor's budget message as reported in the Pittsburgh Post-Gazette (selected issues - 1940-1976).

III.2. Adequacies and Inadequacies of the Data To The Hypotheses

> To estimate the expenditures [in the early 1800's] was not so difficult. These were made in the beginning by the several departments, and then submitted by the secretary of the treasury to Congress. Sometimes they were diminished, sometimes increased. Almost always new expenditures were authorized. When the amount was determined, and also the estimates of receipts which were likely to accrue from existing laws, a foundation was laid for several inquiries. Will the receipts probably be sufficient to pay the expenditures? Shall the revenues be drawn from other sources? or, in case of a deficiency, shall this be bridged by borrowing, or by cutting down expenditures? If not, what new fountains of national income shall be opened? These inquiries were answered in various ways, and too often Congress displayed a painful lack of wisdom and principle in answering them.
>
> - Bolles (1894, Vol.II, p.527)

Problem formulation in budgeting is very difficult to study. There are few trace records or written retrospective accounts of how formulation is accomplished. There are many individuals and several subunits of government organizations who formulate the budget over a several month period. This makes it very difficult to gather protocols on the process. Budget problem formulation, particularly the critical early stages (before the Chief Executive sends instructions to the Departments for the preparation of budget requests), is an informal, frequently secretive process. The participants are not required to prepare or preserve any documents about their reasoning in these early stages. (We do have access to a few such confidential documents for Pittsburgh and other governments). Officials do not usually air the versions of the budget problem prior to the version which

the Chief Executive presents to the Legislature.* The instructions to departments for budget preparation are usually the first semi-public indication of how the budget problem has been formulated.

The formal aspects of the budgetary process are amply recorded as the data collected for this paper attest. But this data has some real limitations for examining our topics. First, the data is aggregate; each observation, including those taken from the Mayor's budget messages, reflect the behavior of several individuals. There is no data on individual behavior. Second, in spite of the numerous observations, they are very low density relative to the processes. And third, the data are "cold" and subject to some interpretative ambiquities. We encountered a number of obvious errors in the records which we corrected. But more troubling are anomalous ovservations that were not obviously errors and for which we could get no firm explanation. Firm explanations for what happened in the handling of bond interest in 1940 or even 1950 are understandably hard to get. Also, there are a few holes in the data from Mayors' budget messages and newspapers. As social science data sets go, this one is extraordinarily clean; but there are still a few problems. We will attempt to point out the most important of these problems as we use the data.

III. 3. Budget Formulation : A Model of the First Trial Balance

Without a record of various formulations of the budget problem in

---

*One interesting variation can be found in Worcester, Massachusetts (see Larkey, 1979) where almost every year the city manager announces the property tax increase that would be required to fund a simple summation of the departments' appropriation requests. This increase is always very large and the practice of announcing the possibility of a very large increase is probably a strategy to condition the community's expectations. Beside the inflated tax increase, the tax increase that results after department requests have been cut appears small.

each of our thirty-seven years, it was not possible to study formulation directly. An alternative approach, the one we adopted, is to build a model that simulates the behavior of City Officials in formulating budget problems.

Based on very extensive interviews with Pittsburgh City Officials with primary responsibility for budget formulation, interviews in which we were essentially told Revenue Source - by - Revenue Source and Expenditure Classification - by - Expenditure Classification how forecasts were prepared, we built a model that looks one year ahead in forecasting revenues and expenditures. Tax Revenue Sources were forecast using a simple moving average process of the following form:

$$\overline{TaxBase\ (K)} = TaxBase\ (K\text{-}1) + (TaxBase\ (K\text{-}1)^* ((Change\ (K\text{-}1) + Change\ (K\text{-}2) + Change\ (K\text{-}3)) / 3))$$

Where

$\overline{TaxBase\ (K)}$ = Forecast TaxBase in Year K

TaxBase (K-1) = Actual TaxBase in Year K-1

Change (K-1) = Rate of Change in Actual TaxBase from Year K-2 to Year K-1

$$\overline{REVYLD\ (K)} = TAXRT\ (K\text{-}1)^* \overline{TAXBASE\ (K)}$$

Where

$\overline{REVYLD\ (K)}$ = Forecast Revenue Yield in Year K

TAXRT (K-1) = TaxRate in Year K-1

All nontax Revenue Sources and most Expenditure Classifications* were

*Where we had reasonable approximations of the real forecasts (E.G., Debt and Interest requirements for year K from the Controller's Report for year K-1) we used their forecasts directly.

forecast using a moving average process identical to the one used above to forecast TaxBase (K). Although the analytic procedures are straightforward, the number of different revenue sources and expenditure classifications makes the model in FORTRAN tediously complex. We will not describe it in any further detail here but we will furnish, upon request, the FORTRAN program.

It is very difficult to validate this model because it is intended to simulate an unbiased First Trial Balance which is never recorded. We believe that the model is plausible and produces problem formulations (including a revenue estimate, an expenditure estimate, and a GAP) which roughly correspond to first trial balances for two reasons.

First, the model was based closely on the description of analytic procedures from the interviews. The model is, however, a more accurate simulation on the revenue side than on the expenditure side. This is because officials have much less discretion on revenues than expenditures. To modify revenues in any significant way, they must change tax or water rates and we know when those changes occur. One source of model error is when the officials have very specific knowledge about a change (e.g., a large new development going on the tax roll) that the model does not have. They have much more latitude on the expenditure side to alter assumptions in formulating the budget problem. Assumptions about wage levels, the number of personnel, pension costs, debt costs and materials costs are obviously important to the problem formulation. Although they do manipulate these assumptions and the assumptions change the budget problem, we view such changes as solution behavior. The officials do indicate that their formulation is a <u>status quo</u> formulation -- if tax rates and levels of personnel and other service inputs remain the same with only "natural"

growth in the revenue bases and the input costs, what will the gap be? There are a few internal documents in years subsequent to our sample period which served as examples of these calculations.

A second reason for some confidence in the plausability of the model is that an indirect check of its validity was possible. We altered the model to forecast for Year K using tax rates for Year K rather than K-1. Under this procedure we would expect the model to produce reasonable forecasts or actual (end-of-year) revenues and expenditures. Table 1 shows the series and differences for revenues and expenditures. Although the percentage mean absolute errors are not as low as we would want if the model's task was to forecast actual revenues and expenditures, the model's task was to simulate problem formulation, the first step in an eighteen month problem-solving sequence that terminates in the "actuals". Indeed, a curious but expected result reported below is that the model exercised with new tax rates provides better forecasts of actuals than the revenue estimates and appropriations in the final budgets.

TABLE 1

COMPARISONS OF PREDICTED AND ACTUAL REVENUES AND EXPENDITURES

CITY OF PITTSBURGH 1940 - 1976

Y.	Revenues (in thousands)			Expenditures (in thousands)		
	Predicted	Actual	Difference	Predicted	Actual	Difference
1940	26657	26222	435	22245	26427	- 4182
1941	25398	24787	611	23754	24695	- 841
1942	23880	23854	26	23908	23801	107
1943	23162	22891	271	24315	23462	853
1944	24835	23689	1146	22054	23617	- 1563
1945	24395	23902	493	22999	23333	- 334
1946	26742	26330	442	25041	26150	- 1109
1947	26656	26866	- 210	27164	26126	1038
1948	31800	31854	- 54	30157	30938	- 781
1949	32110	31935	175	33589	32876	713
1950	33428	33110	318	35963	33737	2226
1951	35829	36238	- 409	35824	34491	1333
19[illegible]	40677	40885	- 218	40121	40515	- 394
1	40089	39976	113	43418	42645	773
19[illegible]4	47810	47735	75	45953	46353	- 438
1955	48110	46724	386	49496	46865	2631
1956	51171	51037	134	49356	47450	1906
1957	50321	49907	414	49663	49526	137
1958	52334	50454	1880	52264	53003	- 739
1959	55346	54031	1315	56550	54240	2310
1960	54646	55148	- 602	57822	54522	3300
1961	59312	56288	3024	56623	57245	- 622
1962	57241	58808	- 1567	59112	58501	611
1963	60859	60093	766	61456	58612	2844
1964	63275	63519	- 244	60723	61174	- 451
1965	64520	64108	412	63254	63644	- 390
[illegible]	68463	68536	- 73	65895	67550	- 1655
19[illegible]7	72491	71699	792	70928	71523	- 595
1968	77779	76687	1112	77204	79966	- 2762
1969	88210	88177	33	88773	90509	- 1736
1970	99602	99411	191	110470	93983	16487
1971	106301	106705	- 404	103685	101591	2094
1972	109542	108427	1115	111248	105514	5734
1973	110578	113626	- 3048	110868	112616	- 1748
1974	110714	108770	1944	122198	107057	15141
1975	106664	109026	- 2352	113712	118678	- 534
1976	125014	124150	864	125555	127634	- 2079

Percent Mean Absolute Error

Predicted Revenues - Actual Revenues = 1.39

Predicted Expenditures - Actual Expenditures = 3.89

III. 4. Evidence on the Pittsburgh Experience

Hypothesis 1: Chief executives present biased formulations of budget problems to legislators and the public. The formulation consists of biased estimates of revenue and expenditure requirements and is calculated to show that the budget problem is worse than the unbiased estimates would show.

Table 2 presents evidence that provides strong support for this hypothesis. To test the hypothesis, we constructed several comparisons between mayor and council estimates and the predicted (from the simulation of the first trial balance) and the actual revenues, expenditures, and budget gap. The mean percent difference and the standard deviation for each comparison was computed and is reported in columns (1) and (2). In addition, the percentage of the years in which the difference was in the expected direction was calculated and is reported in column (3). As we would expect from the observations of prior research, mayors and city councils consistently underestimate the expected revenues (whether measured by both the mean percent error or the percent correct signs), regardless of whether the predicted or the actual revenue figure is used as the baseline. But more importantly, there is evidence that the entire budget problem is misrepresented. Not only are revenues underestimated, but expenditures are overestimated, resulting in budget gaps that are often depicted as worse than they "really" are. The formulation of the budget problem, then, is clearly strategic.

Hypothesis 2: The extent of the bias in the executive's formulation of the budget problem is conditional.

a) Budget problems are presented as worse than they "really" are when there are no strong external pressures (an impending election) to avoid certain solutions to the budget problems (tax increases).

The evidence in Table 3 indicates that the strategic formulation of the budget problem is different in election years, but in ways not completely anti-

TABLE 2

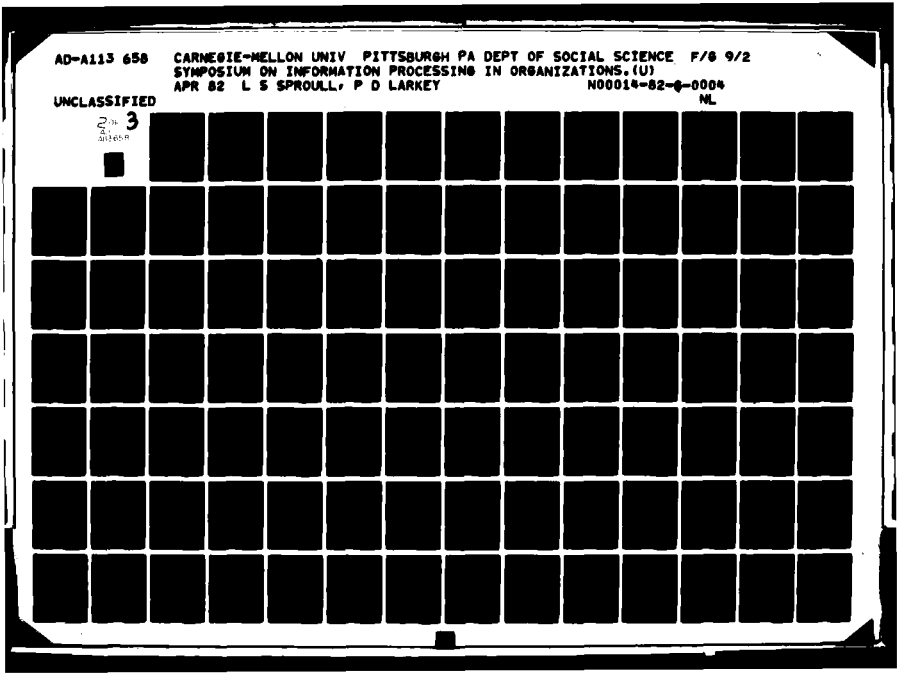
COMPARISONS AMONG ESTIMATED AND ACTUAL REVENUES, EXPENDITURES, AND GAPS

PITTSBURGH 1940 - 1976

Comparison	Mean Percent	Standard Deviation	Percent Correct Signs
Revenues:			
Mayor* - Predicted*	- 3.34	3.45	86.4
Mayor - Predicted	- 1.96	3.12	81.1
Council - Predicted	- 1.99	2.79	86.4
Council - Actual	- 1.26	2.84	73.0
Mayor - Actual	- 1.25	2.90	70.3
Expenditures:			
Mayor* - Predicted	2.36	5.97	64.8
Mayor - Predicted	1.23	4.63	59.4
Council - Predicted	1.63	4.16	62.2
Council - Actual	2.30	2.61	94.6
Mayor - Actual	1.87	2.88	81.1
Budget Gap:			
Mayor* - Predicted*	- 5.39	7.39	78.4
Mayor - Predicted	- 3.22	5.95	64.8

Note: For revenues, the superscript [*] indicates the estimate does not include any proposed revenue changes. For expenditures, the superscript indicates the estimate has not been adjusted as part of the solution to the budget gap. For budget gaps, the superscript indicates that neither revenues or appropriation have been adjusted.

AD-A113 658 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF SOCIAL SCIENCE F/G 9/2
SYMPOSIUM ON INFORMATION PROCESSING IN ORGANIZATIONS.(U)
APR 82 L S SPROULL, P D LARKEY N00014-82-G-0004
UNCLASSIFIED NL



cipated. As part of a more accurate problem formulation, we expected the revenues and expenditure estimates of the mayors would be more accurate (i.e., less underestimation of revenue and less overestimation of expenditure requirements). The evidence presented in Table 3, however, suggests that our expectation is only partially correct. Revenues are more accurately estimated, but expenditure requirements are not. In election years, the mayors' estimates of expenditures are, on average, larger. The larger averages are consistent with a strategy in which mayors introduce new programs and services and expand old ones in election years. However, generally, the adjustments on the expenditure side are smaller than the adjustments on the revenue side, meaning that the overall formulation of the budget problem is more accurate in election years.

In Table 3, however, there seems to be one exception to this conclusion. The difference in the predicted and the mayor's budget gap estimate, calculated prior to any proposed revenue changes or expenditure adjustments (i.e. Mayor* - Predicted*) is considerably larger in election years than non-election years. Yet, this larger difference should not be interpreted as an exception. Instead, it reflects the existence of two confounding factors: 1) the formulation of budgets in two years by mayors who were not seeking reelection and, therefore, lacked the usual incentives to present more accurate formulations; and 2) the impending use of a budget gap solution that is used primarily in election years and gives the mayors the luxury of not having to rein in their estimates of expenditure requirements.

The solution is the refunding of maturing bonds. Typically, maturing bonds are retired using general fund revenues. But sometimes mayors refund maturing bonds (by selling new bonds to refinance them) and general fund revenues are no longer needed. Mayors propose refunding because it is a painless solution for

TABLE 3

COMPARISONS OF ESTIMATED AND ACTUAL REVENUES, EXPENDITURES AND BUDGET GAPS IN ELECTION AND NON-ELECTION YEARS

CITY OF PITTSBURGH 1940 - 1976

Comparisons	Election Year (n = 18)			Non-Election Year (n = 19)		
	Mean Percent	Standard Deviation	% Correct Signs	Mean Percent	Standard Deviation	% Correct Signs
Revenues:						
Mayor* - Predicted*	- 3.32	3.45	88.9	- 3.41	3.45	84.2
Mayor - Predicted	- 1.66	3.11	83.3	- 2.24	3.12	78.9
Council-Predicted	- 1.59	3.21	83.3	- 2.37	2.27	89.5
Council - Actual	- .82	3.65	55.6	- 1.69	1.66	89.5
Ma - Actual	- .91	3.30	61.1	- 1.57	2.60	78.9
Expenditures:						
Mayor* - Predicted	3.16	5.26	67.7	1.60	3.65	63.2
Mayor - Predicted	1.45	4.28	55.6	1.01	4.93	63.2
Council - Predicted	1.60	3.65	55.5	1.67	4.59	68.4
Council - Actual	2.46	1.59	94.3	2.14	3.30	94.7
Mayor - Actual	2.31	2.24	83.3	1.46	3.33	79.0
Bu..c Gap:						
Mayor* - Predicted*	- 6.61	6.54	83.3	- 4.72	8.19	73.7
Mayor - Predicted	- 3.04	5.01	61.1	- 3.36	6.84	68.4

Note: For revenues, the superscript [*] indicates the estimate does not include any proposed revenue changes. For expenditures, the superscript indicates the estimate has not been adjusted as part of the solution to the budget gap. For budget gap, the superscript indicates that neither revenues or appropriation have been adjusted.

reducing the budget gap; it serves to reduce expenditure requirements in a given year by moving them to later years. Because it is painless, refunding is a popular solution in election years. It allows mayors to continue to increase expenditures without having to raise revenues. Over the period of 1940 to 1976, refunding was proposed by mayors five times as a solution to the budget gap and four of those proposals occurred in election years. In each case, the amount of bonds refunded totaled more than ten percent of that year's total expenditures. Therefore, in election years in which refunding is proposed, mayors anticipate having substantially more revenue available for non-debt spending, leaving them few reasons to estimate expenditures lower than they do in non-election years.

Table 4 presents some additional evidence for the effects of elections on problem formulation. First, the size of the budget gap reported by the mayor is larger in election years than non-election years. This is especially true when the election year budgets of "lame duck" mayors (mayors who have decided not to run for reelection) are deleted from the analyses. Second, the solution proposed by the mayor to close the budget gap is different in election years than non-election years. In election years, mayors proposed that about seventy-five percent of the budget gap be closed by increasing taxes; in non-election years, they typically proposed that almost the entire gap be closed by raising taxes. These results when combined with those in Table 3 offer strong support for hypothesis 2a.

Hypothesis 2: The extent of the bias in the executive's formulation of the budget problem is conditional.

b) Budget problems are presented as worse than they "really" are when tax changes (either increases or reductions) are proposed than when no change (especially in those years prior to a tax increase) is proposed. Chief executives, in other words, extract (in years

TABLE 4

THE EFFECTS OF ELECTIONS ON THE BUDGET GAP AND SOLUTIONS

REPORTED BY THE MAYOR IN YEARS OF PROPOSED TAX INCREASES

CITY OF PITTSBURGH 1940 - 1976

	Election Year (n = 6)	Non Election Year (n = 12)
Size of the reported budget gap (as a percentage of the total proposed expenditures)	6.1 (5.0)	7.4
Percentage of the budget gap eliminated by increasing taxes	76.3 (75.0)	96.3

Note: Cell entries in parentheses exclude 1959 and 1969. In each year, the budget problem and solution were formulated by mayors who were not seeking reelection.

of tax increases) and retain (in years of tax reductions) resources greater than they "really" need.

Table 5 presents evidence that strongly supports this hypothesis. Whether measured by the mean percent difference or the percentage of the differences that are in the hypothesized direction, there is a much stronger bias in the formulation of budget problems in years of tax increases and reductions compared to years in which tax rates remain the same. Revenues are consistently underestimated more frequently and to a greater degree; expenditure requirements are consistently overestimated more often and to a larger extent; and, as a result, the reported budget gap is consistently overestimated more frequently and to a greater extent. The evidence, therefore, is clear: mayors take advantage of the occasion of a tax increase to ask for more financial resources than they "really" need and mayors are careful in years of tax reductions to retain more resources than the city's current financial situation demands.

Hypothesis 2: The extent of the bias in the executive's formulation of the budget problem is conditional.

c) Budget problems are presented worse than they "really" are when expected revenues exceed or approximate expected expenditures. Budget formulation becomes more accurate as the size of projected deficit increases.

In order to test this hypothesis, we constructed an index of "fiscal pressure."

$$\text{Fiscal Pressure}_{(t)} = \left[\widehat{\text{Revenues}}^{*}_{(t)} - \widehat{\text{Expenditures}}_{(t)}\right] / \widehat{\text{Expenditures}}_{(t)}$$

where $\widehat{\text{Revenues}}^{*}_{(t)}$ = predicted revenues for year$_{(t)}$, not adjusted for any proposed tax changes in t.

$\widehat{\text{Expenditures}}_{(t)}$ = predicted expenditures$_{(t)}$

Fiscal pressure values are negative in years of projected deficits; positive in

TABLE 5

COMPARISONS OF ESTIMATED AND ACTUAL REVENUES, EXPENDITURES AND BUDGET GAPS BY NATURE OF THE PROPOSED TAX ACTION

CITY OF PITTSBURGH 1940 - 1976

	Tax Increase (n = 18)			No Change (n = 12)			Tax Reduction (n = 7)		
	Mean Percent	Standard Deviation	% Correct Signs	Mean Percent	Standard Deviation	% Correct Signs	Mean Percent	Standard Deviation	% Correct Signs
Revenues:									
Mayor* - Predicted*	- 3.79	3.33	88.9	- 2.32	3.88	75.0	- 4.09	2.41	100.0
Mayor - Predicted	- 2.13	3.11	83.3	- 1.57	2.83	75.0	- 2.21	3.56	85.7
Council - Predicted	- 2.10	2.30	94.4	- 1.88	2.71	75.0	- 1.90	3.88	85.7
Council - Actual	- 1.09	2.51	77.8	- 1.00	2.44	66.7	- 2.19	3.90	71.4
Mayor - Actual	- 1.13	2.95	61.1	- .69	2.50	75.0	- 2.50	3.44	85.7
Expenditures:									
Mayor* - Predicted	3.31	6.32	72.2	.97	6.92	41.7	2.31	3.72	85.7
Mayor - Predicted	2.10	4.93	66.6	- .82	3.92	33.3	2.48	3.73	85.7
Council - Predicted	2.83	4.46	72.2	- .82	3.07	33.3	2.78	2.94	85.7
Council - Actual	2.07	3.04	94.4	2.24	2.20	92.7	2.96	1.83	100.0
Mayor - Actual	1.33	3.04	77.8	2.22	2.04	92.7	2.67	3.39	71.4
Budget Gap:									
Mayor* - Predicted*	- 6.67	7.84	88.9	- 3.13	7.62	58.3	- 6.14	4.86	100.0
Mayor - Predicted	- 4.32	6.47	72.2	- .79	5.44	50.0	- 4.48	4.31	85.7

Note: For revenues, the superscript [*] indicates the estimate does not include any proposed revenue changes. For expenditures, the superscript indicates the estimate has not been adjusted as part of the solution to the budget gap. For budget gaps, the superscript indicates that neither revenues or appropriations have been adjusted.

years of projected surpluses. When predicted revenues and expenditures are equal, fiscal pressure is 0. Figure 2 shows the variability in fiscal pressure for the City of Pittsburgh over the thirty-seven year period under study.

According to our hypothesis, the formulation of the budget problem by the mayor becomes more accurate as fiscal pressure grows (i.e. its value becomes increasingly negative). Fiscal pressure, therefore, should be negatively correlated with the difference in the size of budget gaps estimated by the mayor and predicted by the simulation: the greater the fiscal pressure, the smaller the difference in the size of the gap. In addition, fiscal pressure should be negatively correlated with differences in the mayor's estimate of revenues and the predicted revenues: the greater the fiscal pressure, the smaller the underestimation of revenues by the mayor. Finally, fiscal pressure should be positively correlated with the difference in the mayor's estimate of expenditures and the predicted estimates: the greater the fiscal pressure, the smaller the overestimation of expenditures by the mayors. Column 1 of Table 6 reports correlation coefficients that are consistent with the above expectations.

The only exceptions are two expenditure comparisons: council-actual and mayor-actual. The negative correlations suggest an important insight into problem formulation by mayors: despite a pronounced tendency to overestimate expenditure requirements - a tendency that is only moderated by the level of fiscal pressure - mayors' actual expenditures more closely reflect the financial condition of the city. Mayors tighten up spending policy in years when fiscal pressure is high. This, in turn, leads to greater differences between stated and actual spending policies in those years and, hence, a negative correlation. Mayors may give untrustworthy reports, but they are not irresponsible.

The relationship between fiscal pressure and problem formulation was also investigated in years in which elections were held and in years when different

FIGURE 2

CITY OF PITTSBURGH – FISCAL PRESSURE 1940 – 1976

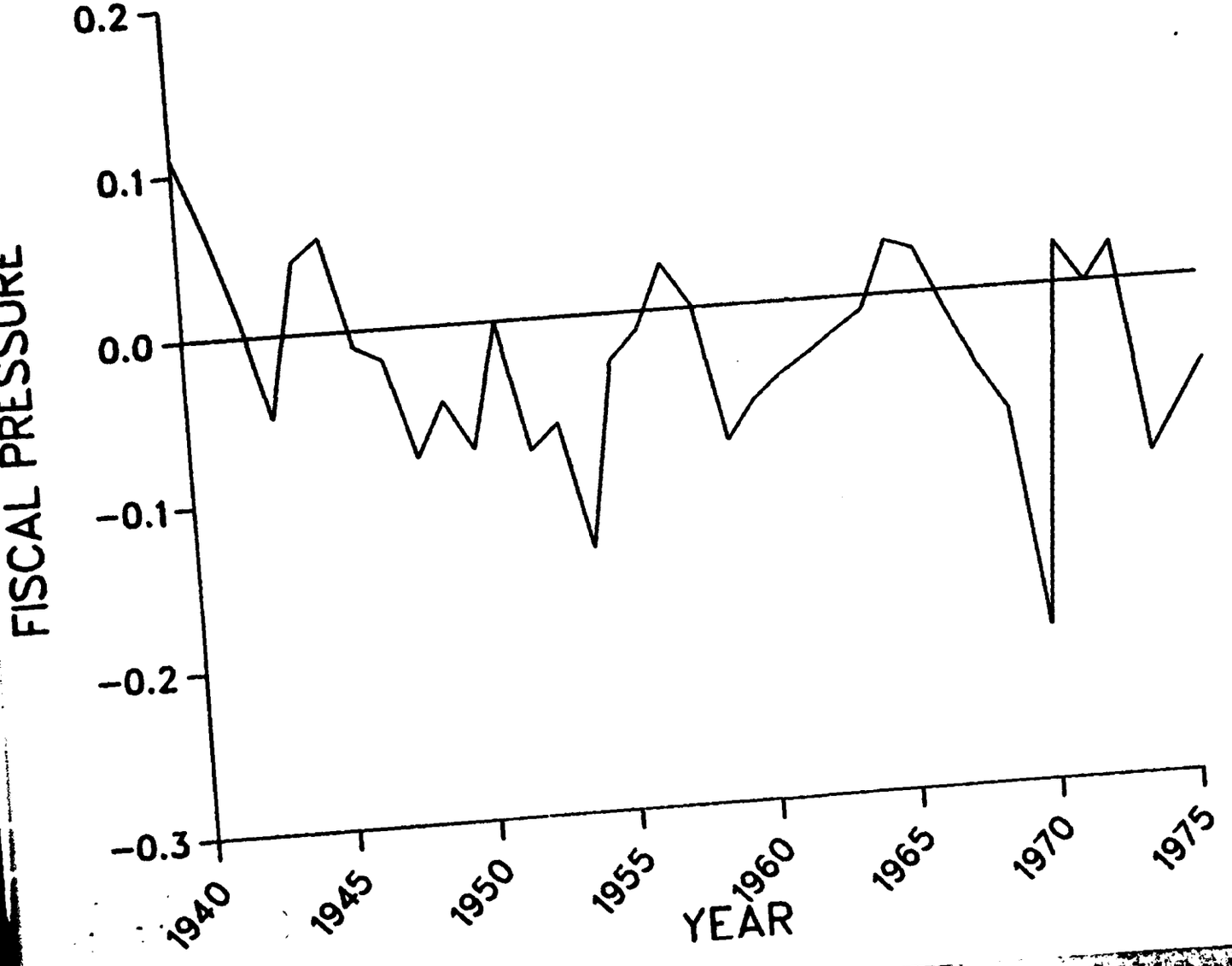


TABLE 6

FISCAL PRESSURE AND PROBLEM FORMULATION

CITY OF PITTSBURGH 1940 - 1976

Comparisons	All Years	Election Year: Yes (n = 18)	Election Year: No (n = 19)	Nature of Proposed Tax Action: Increase (n = 18)	Nature of Proposed Tax Action: No Change (n = 12)	Nature of Proposed Tax Action: Tax Reduction (n = 7)
Revenues:						
Mayor* - Predicted*	- .53	- .55	- .56	- .60	- .49	- .61
Mayor - Predicted	- .58	- .54	- .69	- .67	- .47	- .97
Council - Predicted	- .37	- .53	- .39	- .49	- .15	- .72
Cou[illegible] - Actual	- .28	- .49	- .25	- .29	.0003	- .65
Mayor - Actual	- .53	- .55	- .64	- .57	- .38	- .70
Expenditures:						
Mayor* - Predicted	.67	.55	.72	.77	.83	.01
Mayor - Predicted	.58	.40	.70	.66	.86	.06
Council - Predicted	.50	.25	.65	.62	.84	- .05
Council - Actual	- .45	- .02	- .61	- .55	- .55	- .07
May - Actual	- .21	.33	- .47	- .39	- .21	.14
Budget Gap:						
Mayor* - Predicted*	- .74	- .71	- .76	- .82	- .89	- .36
Mayor - Predicted	- .76	- .67	- .84	- .85	- .86	- .65

Note: Cell entries are Pearson correlations between fiscal pressure and the various comparisons of revenues, expenditures, and gaps indicated above.

kinds of tax actions were proposed. Following hypothesis 2a, we would expect the relationship between fiscal pressure and problem formulation to be weaker in election years than non-election years. When faced with an impending election, we would expect that mayors would moderate their strategic behavior, thereby reducing the extent to which they overestimate expenditures and underestimate revenues when fiscal pressure is low. However, in light of the evidence presented in Table 3, we would further expect that the degree of moderation would be small. The differences in correlation in election versus non-election years should be slight.

Columns 2 and 3 of Table 6 present correlation coefficients that are generally consistent with our expectations. The expected difference is most clearly visible in overall problem formulation: the difference in the mayor's estimate of the budget gap and the predicted gap. In the case of expenditures, the evidence is consistent but with an interesting twist. For the council-actual and the mayor-actual comparisons, the correlations are non-existent and weakly positive in election years and moderately negative in non-election years. The differences suggest that, despite the overall responsible spending policies noted above, mayors are reluctant to tighten down actual spending in election years, even when the city's financial condition warrants it. Elections may improve the trustworthiness of verbal reports about the budget problem, but they do not seem to encourage responsible spending policies. Finally, on the revenue side, there is some evidence of the expected difference for the mayors' estimates, but none for council. In fact, the relationship for council is inexplicably weaker in non-election years than election years.

The relationship between fiscal pressure and problem formulation under different kinds of proposed tax action is presented in Columns 4, 5, and 6 of Table 6. Here, the relevance of hypothesis 2b as a basis for interpreting the evidence is not clear. Fiscal pressure is likely to be greatest in years of tax increases

and least in years of tax reductions. Thus, the range of variation in fiscal pressure is likely to be small in years of either kind of tax action. In tax reduction years, the restricted range of variation is especially severe: there are only projected surpluses or at worst small deficits. Real fiscal pressure does not exist. So it is unclear what the relationship between fiscal pressure and problem formulation might be. In tax increase years, however, the expected relationship should be strongly negative. Mayors are faced with considerable fiscal pressure and they have made the decision to increase taxes. In fact, about the only way that their formulations would not accurately reflect the financial needs of the city is that they might be overly pessimistic in hopes of extracting additional financial resources. As for those years in which taxes remain the same, the expected relationship should also be strongly negative. In these years, when fiscal pressure is high, the mayor's formulation of the budget problem must be more accurate: there are few slack resources available and the possibility of seeking a tax increase has been ruled out. The correlation coefficients reported in Columns 4, 5, and 6 are generally consistent with these expectations.

Hypothesis 3: Chief executives present biased public justifications of the formation of the budget problem. The bias occurs in the justifications of proposed increases in expenditures and takes two forms.

a) Chief executives single out for attention only a subset of the proposed increases in expenditures for specific programs, typically those viewed as highly desirable by the legislature and the public (e.g., public safety and parks and recreation).

b) Chief executives overstate the proportion of the total increase in expenditures that are "really" due to increases in desirable expenditures.

There is little evidence for this hypothesis. Analyses of the mayor's budget messages indicate that the hypothesized behaviors are not prominent explanatory strategies. Instead, what emerges from the analyses as the dominant presentation strategy is something quite different: Officials explain the budget formulation in ways that absolve them of most of the responsibility for it. Their basic tactic is

to portray the budget problem as largely due to factors beyond their control. That means their presentations emphasize, above all, bad news: inflation is driving city expenditures up at a rate faster than revenues are growing, union arbitration awards are having a ruinous effect on city finances, revenues projections are down because of an area-wide recession, and so on. And any good news that is presented is qualified by the bad: gains in revenues are presented as offset by mandated expenditure increases or unexpectedly large surpluses from the prior year are countered with views of unexpected revenue losses.

As an illustration of the bad news and not so good news phenomenon, consider the introductory paragraphs of the mayor's message accompanying the 1968 budget:

> This General Fund Budget for 1968, which I have the honor to present to you today, is one of the most troublesome I have ever submitted in my tenure as Chief Executive of this city.
>
> It is troublesome because it is built on a seeming contradiction. On the one hand, Pittsburgh fortunately still finds itself in the midst of a sustained economic surge with relatively low unemployment throughout the labor market area. At the same time, the City's resources are simply inadequate to meet our overall municipal obligations from year to year.
>
> We will finish the year with a healthy net surplus of $4,205,000 available for the 1968 Budget. This reflects generally good revenue collections in 1967 exceeding the original estimates by approximately $1,700,000 and cancellations from all Code Accounts amounting to $2,505,000.
>
> Yet, the strong year-end balance is offset by several major adverse developments during the past 12 months.
>
> 1. Pittsburgh has been sharply set back by the recent court decisions reducing Downtown valuations which required the City Treasurer to make emergency unanticipated refunds of $570,000 this year and which will apparently mean even greater rebates in 1968. Indeed, the effect of these reductions in property values may run into millions of dollars in the years ahead.
>
> 2. The City's stepped up pace of physical improvements, coupled with the high cost of borrowing in the present tight money market, is adding nearly $1-Million to the City's Debt Service obligations.
>
> 3. The Nation is caught in an inflationary spiral which produces higher prices for every service the City performs and for every item the City buys from Rock Salt to Rubber Tires.

> We are in no better position to fend off inflation than the individuals and families who make up the City population and who pay the City Taxes.
>
> There is no foreseeable sign of abatement in this inflationary pattern; at the same time there is no sign of abatement in the public demand for improved and expanded services.
>
> We start the New Year with mandatory increases - by definition expenses over which we have no choice - amounting to nearly $2-Million.

The printed message continues in this fashion. As it does, its underlying message to legislators and citizens alike becomes unmistakably clear: "You may not like the problem I am presenting, nor the solutions I am proposing. I don't like them either. But my hands are tied; I have no other choice. If you dislike it, that's fine. But don't blame me. I'm not responsible for it."

Two sets of analyses of the mayor's messages are presented in support of the above argument. The first set, reported in Table 7, summarizes the ways in which proposed expenditure increases are explained by mayors. These explanations have two parts. The first part consists of presenting the total proposed increase in expenditures, using the prior year's total appropriations figure as the baseline. For the period 1947 to 1976 (the years in which budget messages are available), penditure increases were presented in this way for twenty-four years.* The second part involves the characterization of the expenditure increase. Mayors depict the increase as falling almost exclusively into two categories: mandated and obligatory increases. Mandatory expenditure increases are completely beyond the control of officials. They include increases due to factors such as increased costs of supplies, equipment and utility service, union arbitration settlements, or state laws reducing the work week of police and fire personnel, thereby requiring the addition of more personnel to maintain the same level of service. Obligatory

---

*In two other years, there was a net reduction in proposed expenditures over the prior year's appropriation. In the remaining four years, a total increase over the prior year's appropriation was not presented, although individual increases were reported.

expenditure increases are ultimately discretionary, but they are presented as increases the city has a strong obligation to approve. They include cost of living adjustments to wage and pension benefit, salary upgrading due to job reclassification plans, matching of the prevailing community/union wage rates, and the institution of group health and life insurance plans to keep pace with the standards set by the private sector.

Part I of Table 7 presents the mean percentages of the total expenditure increases attributed to each of these categories. The results are also broken down by the nature of the proposed tax action in a given year (i.e., whether officials proposed a tax increase, no change, or a tax reduction).. On average, nearly sixty percent of any expenditure increase is presented as mandatory while nother thirty-six percent is portrayed as obligatory. Overall, an average of ninety-three percent of the increase is explained as largely beyond the control of the mayor. When broken down by the nature of the proposed tax action, the pattern continues with some variation. Most important is that mayors claim to have the least discretion over proposed expenditure increases in years of tax increases and the most discretion in years of tax reductions. In other words, mayors say they have the least control over rising expenditures in the years they proposed the most unpopular solutions to the budget problem.

Part I of Table 7, however, understates the extent to which mayors try to categorize rising expenditures as mandatory. In years in which the percentage of the total increase is mandated, mayors calculate two increases: the total increase report in Part I and an increase that excludes obligatory wage and benefit increases. This second increase is the "basic budget" increase. Part II of Table 7 demonstrates the effect that this dual characterization of budget increases has on the explanation of rising expenditures. First, in years in which a basic budget is not presented, an average of seventy-five percent of the total increase is categorized as mandatory compared to only fifty-seven percent

TABLE 7

CHARACTERIZATIONS OF PROPOSED EXPENDITURE INCREASES WHEN APPROPRIATIONS$_{(t-1)}$ SERVES AS THE BASELINE*

SELECTED YEARS 1947 - 1976 CITY OF PITTSBURGH

PART I:

Category	Nature of Tax Action: Tax Reduction (n = 15)	No Change (n = 5)	Tax Reduction (n = 4)	All Years (n = 24)
bligatory Increases	42.9	29.5	30.4	36.1
landated Increases	55.3	64.7	49.0	57.2
	98.2	94.2	79.4	93.3

PART II:

Category	"Basic Budget" Presented (n = 10): "Basic Budget" Increase	"Basic Budget" Presented (n = 10): Total Increase	Basic Budget Not Presented (n = 14): Total Increase
bligatory Increases	2.0	59.5	19.4
landated Increases	90.6	33.2	74.3
	92.6	92.7	93.7

*Cell entries are the percentages of the increase attributed to a particular category.

when all years are considered. Second, in those years when a basic budget is presented, an average of over ninety percent of the basic budget increase is categorized as mandatory compared to only thirty-three percent for the total increase. In effect, the construction of a basic budget allows mayors to highlight "how tightly their hands are tied." Attention is no longer focused solely on an increase of which only a moderate percentage can be characterized as mandatory. Instead, attention is also directed toward a budget and increases over the prior year's appropriations that are almost completely mandated. Mayors, as a result, can place greater emphasis on the extent to which rising expenditures are beyond their control.

The second set of analyses summarize the reasons given by mayors in their explanations of the yearly budget problem. The analyses concern the reasons presented in the mayor's budget messages for the period 1947-1976. In each message, the reasons given for the form of the budget problem were coded, categorized, and counted. Admittedly, a simple count of the reasons given has its shortcomings. The most important one is that it does not capture well the relative emphasis given to different reasons. That emphasis is not only a function of frequency of mention, but also the linguistic qualifiers employed and the structure of the overall explanation. Still, a simple count of the reasons does capture a great deal of the information presented in the messages. Accordingly, it is a useful first "cut" at the problem and is presented in that spirit. The categories of reasons used in the analyses are presented in Chart 1.

Three indices were constructed in order to summarize the number of times that reasons from different categories appeared in the explanations offered over the thirty year period.

Index (1) is the proportion of years in which one or more reasons within a category were cited.

Index (2) is the mean number of times the reasons within a category were cited each year.

Index (3) is the mean number of times the reasons within a category were cited each year, adjusted for the number of reasons within each category.

CHART 1

CATEGORIES OF REASONS GIVEN FOR THE NATURE OF THE BUDGET PROBLEM

MAYOR'S BUDGET MESSAGES - 1947 - 1976

CITY OF PITTSBURGH

1) Uncontrollable revenue losses

- Loss of Federal/State funds
- Decline in real estate valuations
- Tax yields lower than usual (typically because of recession)
- Loss of miscellaneous funds/sales of City property

2) Mandatory expenditure increase

- Inflation
- State/Federal mandated expenditure increases
- Mandatory wage increases
- Mandatory benefits/pension increases
- Debt/sinking fund requirements
- Court ordered tax refunds

3) Surplus lower than expected

4) Unfavorable Federal/State/County policies

- Insufficient assumption of city services
- Inadequate transfer payments
- Unrealistically low assessments of real estate values

5) Obligatory wage increases granted

- Cost of living/prevailing community and union wage adjustments
- Special adjustments (e.g., job reclassifications, pension adjustments, initiation and improvement of group life/medical insurance plans)

6) Obligatory wage increases denied

7) "Necessary" program increases

- Public safety
- Parks and recreation
- Health
- Housing
- Community grants
- Public works, street lighting, traffic, planning

8) Program Reductions

- Parks and recreation
- Community grants
- Social services
- Public Works
- Miscellaneous free services

9) Savings due to "efficiency" measures

- Elimination of unfilled/unnecessary/"political" positions
- Modernization/conservation
- Increased utilization of existing work force, especially transfer of contract work to city crews

10) Revenue Increases under existing tax rates

- Increase in Federal/State aid
- Increase in real estate valuations
- Tax yields higher than usual
- Sale of City property

11) Surplus higher than expected

12) Reductions in departmental requests

- Requests reduced
- Further reductions impossible/ill advised

These indices are reported in Table 8, broken down by the nature of the proposed tax action.

Two observations are warranted in light of results reported in this table. First, as expected, mayors emphasize that much of the budget problem is due to factors beyond their control. Examination of the three indices reveals numerous references to obligatory wage increases and uncontrollable factors, especially in those years when a tax reduction is not proposed. In fact, references to uncontrollable factors are most numerous in those years when mayors proposed unpopular solutions to the budget problem (tax increases). And the category of reason most relied upon is (not surprisingly) mandatory expenditure increases. Second, despite the reference to obligatory and uncontrollable factors, mayors do admit that some proportion of the total increase in expenditure is due to "necessary" program increases. As was shown in Table 7, the proportion is not large, typically about ten percent of the total expenditure increase. Mayors, therefore, talk about specific program increases as only marginal additions to an overall expenditure increase that is otherwise beyond their control. Perhaps, that is why references to program increases are remarkably high regardless of the nature of the proposed tax action. Since program increases comprise such a small proportion of the total increase, mayors simply do not consider making major adjustments in their discretionary spending plans.

There is the possibility, however, that the emphasis by mayors on uncontrollable factors never gets communicated beyond council chambers. Most citizens, after all, never read or hear a mayor's budget message. What they do hear is largely presented by the various media covering the budget presentation to council. Here the possibility for distortion exists. Journalists may convey a different explanation of the budget formulation than the mayor presented. They may, for example, emphasize the increases in programatic expenditures and ignore the mayor's emphasis on uncontrollable factors.

TABLE 8

THE RELATIONSHIP BETWEEN CATEGORIES OF REASONS AND THE NATURE OF THE PROPOSED TAX ACTION

MAYOR'S BUDGET MESSAGES 1947 - 1976

CITY OF PITTSBURGH

Categories	Index (1)			Index (2)			Index (3)		
	Tax Increase	No Change	Tax Reduction	Tax Increase	No Change	Tax Reduction	Tax Increase	No Change	Tax Reduction
Uncontrollable Factors:									
Uncontrollable Revenue Losses	.63	.22	0	1.06	.22	0	.27	.06	0
Mandatory Expenditure Increases	1.00	.89	.40	3.06	2.20	2.00	.51	.37	.33
Surpluses Lower than Expected	.25	.22	0	.25	.22	0	.25	.22	0
Unfavorable Federal, State County Policies	.88	.33	.40	1.25	.44	.40	.42	.15	.13
Obligatory Wage Increases Granted	.81	.56	.80	.81	.56	.80	.81	.56	.80
Obligatory Wage Increases Denied	.13	.22	0	.13	.22	0	.13	.22	0
"Necessary" Program Increases	.88	.78	1.00	2.10	2.60	3.20	.26	.33	.40
Program Reductions	.25	.22	.20	.31	.33	.20	.06	.07	.04
Revenue Increases Under Existing Tax Rates	0	.11	.80	0	.11	1.00	0	.02	.20
Surpluses Larger than Expected	.13	.33	1.00	.13	.33	1.00	.13	.33	1.00
Savings due to "Efficiency" Measures	.25	.56	.20	.25	1.00	.20	.08	.33	.07
Reductions in Departmental Requests	.69	.78	.20	1.00	.45	.20	.50	.45	.10

Note: Tax increases were proposed in sixteen years; tax reductions were proposed in five years. No changes in taxes were proposed in nine years.

Index (1) = number of years in which one or more reasons within a category were cited/number of years in which the proposed tax action occurred.

Index (2) = number of citations of reasons within a category/number of years in which the proposed tax action occurred.

Index (3) = Index (2)/number of reasons within a category.

Data relevant to the question of distortion are presented in Table 9. The data reported are taken from accounts of the mayor's budget messages as reported in the Pittsburgh Post Gazette. The overall smaller size of the values of the three indices (compared to those in Table 8) simply reflect the fact that fewer reasons of explanation are reported in the news account than the mayor actually presented. The omission of reasons, however, is not completely random. Uncontrollable factors still dominate explanations in those years when no tax reduction is proposed, but the category of obligatory wage increases is virtually ignored. As a result, the explanations that are conveyed to citizens place even a greater relative emphasis on variables outside the control of public officials. The distortion works to the advantage of the mayor: the "bad news" receives relatively greater emphasis and the mayor appears to be even less responsible for the budget than his own explanation indicates.

IV. Discussion

> One may solve one's problems not only by getting what one wants but also by wanting what one gets.
>
> - Ralph Barton Perry*

We have discovered in Pittsburgh that one may also solve one's budget problems and shirk responsibility for both problems and solutions by presenting the problem as overconstrained and beyond one's control. The problems are formulated and presented as if there is only one feasible solution and no discretionary choice. And of course, with no choice there is no exercise of preferences, ergo, nothing in the problem-solving behavior which any reasonable person could

---

*Paraphrased in Reitman (1967), p.308.

TABLE 9

THE RELATIONSHIP BETWEEN CATEGORIES OF REASONS AND THE NATURE OF THE PROPOSED TAX ACTION

MAYOR'S BUDGET MESSAGES AS REPORTED IN THE PITTSBURGH POST-GAZETTE

1947 - 1976

Categories	Index (1)			Index (2)			Index (3)		
	Tax Increase	No Change	Tax Reduction	Tax Increase	No Change	Tax Reduction	Tax Increase	No Change	Tax Reduction
Uncontrollable Factors:									
Uncontrollable Revenue Losses	.25	.22	0	.31	.22	0	.08	.06	0
Mandatory Expenditure Increases	.50	.44	0	1.50	.56	0	.25	.09	0
Surpluses Lower than Expected	.19	.22	0	.19	.22	0	.19	.22	0
Unfavorable Federal, State, County Policies	.44	.22	.40	.50	.22	.40	.17	.07	.13
Obligatory Wage Increases Granted	.06	.11	0	.06	.11	0	.06	.11	0
Oblagatory Wage Increases Denied	0	.22	0	0	.22	0	0	.22	0
"Necessary" Program Increases	.25	.33	.60	.25	.44	.60	.03	.06	.08
Program Reductions	.13	.22	0	.25	.33	0	.05	.07	0
Revenue Increases Under Existing Tax Rates	0	0	.80	0	0	.80	0	0	.16
Surpluses Larger than Expected	.13	.33	1.00	.13	.33	1.00	.13	.33	1.00
Savings due to "Efficiency" Measures	.13	.44	.20	.13	.44	.20	.04	.15	.07
Reductions in Departmental Requests	.63	.22	0	.69	.33	0	.35	.17	0

Note: Tax increases were proposed in sixteen years; tax reductions were proposed in five years. No changes in taxes were proposed in nine years.

Index (1) = number of years in which one or more reasons within a category were cited/number of years in which the proposed tax action occurred.

Index (2) = number of citations of reasons within a category/number of years in which the proposed tax action occurred.

Index (3) = Index (2)/number of reasons within a category.

find objectionable. This persistent ploy, therefore, allows mayors to present budget problems and solutions while evading the responsibility for the formulation.

In addition, we have found that Pittsburgh mayors routinely misrepresent the budget problem, overstating its severity. They misrepresent the problem by underestimating revenues and overstating expenditures, creating a gap to be closed that is larger in the case of the deficit and smaller in the case of a projected surplus than unbiased estimates of gaps. Some of the ways in which underestimation is accomplished are arcane. For example, Pittsburgh revenue estimations are often low because budget-makers use pessimistic tax collection rates. This finding of misrepresentation is consistent with the finding in the budgeting literature that officials in many types of government underestimate revenues.

There are several reasons why officials overstate budget problems. First, the more difficult and threatening the problem, the more highly motivated people are to search for solutions. Under conditions of budget stringency, employees are more apt to be cost conscious and to generate money saving solutions. In addition to simply announcing the difficulty of the budget problem, budget offices must be the picture of austerity if they are to be credible to the rest of the organization. Budget offices are perpetually underfunded and undermanned because of the need to "practice what they preach."

Second, overstating the budget problem loosens constraints, essentially expanding the solution space. If you can make people believe that the situation is desperate, they expect and will be more accepting of desperate solutions. In Pittsburgh, every second or third year over thirty-seven years was "the tightest year the city had ever faced." Since "tightness" is relative to prior years, officials may have been telling the truth, but we find the frequency of such statements a bit suspicious; some of them must be strategic misrepresentation.

The fiscal pressure measure was not increasing monotonically over the 37 year period.

Third, and finally, misrepresenting the budget problem hedges against two types of uncertainties: environmental uncertainties and strategic uncertainties. Environmental uncertainties include such acts of God or any identifiable scapegoats such as economic downturns that reduce revenues or severe winters that increase costs. Strategic uncertainties are uncertainties about how others in the process will respond to the initial problem formulation and proposed solutions. If city council is in the habit of adjusting the revenue estimate and appropriations upwards, budget officials assume such behavior in their estimates.

A more reasonable way to hedge against these uncertainties would be to budget a contingency account. This has been done periodically in Pittsburgh. But when the fiscal pressure is great, abolishing the contingency account is often the first step in a solution. Also, city councils are reluctant to give mayors discretionary expenditures, particularly when the mayor and council are competing rather than cooperating.

Misrepresenting the budget problem appears to be a necessary part of the role of those with primary responsibility for budgeting. It is not a necessarily self-serving or morally repugnant act. From the role perspectives, it is one of many acts that protect politicians and the heads of service departments from themselves; it is an act that makes "financial responsibility" an operative constraint. Absent the constraint, politicians and heads of service departments would almost always opt for lower taxes and greater expenditures. This constraint, of course, does not guarantee financial responsibility beyond balancing the budget.

It is clear from prior research and the work reported here that city offi-

cials do not formulate their budget problems as the normative, utility maximizing theories would have them formulate the problems. The formulation of budget problems as gaps between projected revenues and expenditures is ubiquitious in government at all levels. The formulation is also found in business firms and households; if these units were required to represent publicly a balanced budget prior to the start of their fiscal years, we would expect to see more overt use of the gap formulation. The appeal of the gap formulation is that it is well structured problem and it connects with many other problems in the organization: the public desires for lower taxes and greater services enter the budget problems as constraints on solution behavior.

Context is very important to problem formulation, solution and justification. We have examined only a few elements of context here, focusing on factors such as elections and fiscal pressure which we hypothesize would be underlying important elements of contexts. Although some of these contextual effects were clear, there are many idiosyncratic elements of contexts which muddy the analyses. We have noted the importance of such idiosyncratic elements as lameduck administrations, and the refunding maturing bonds in our interpretations of the data.

The importance of contexts is not strictly an organization phenomenon. For example, Tversky and Kahneman (1981) note the importance of contexts to individuals in decision framing; one compelling example they give is that of horse players making bets on the last race of the day that would be unacceptable bets on earlier races. The behavior would appear to violate the economists dictum "to ignore some costs" as the horse players with losses attempt to "get even" by taking long odds.

Much work remains before we will have reasonable, useful theories of budgetary behavior in contexts. There are other contextual factors that we hy-

pothesized as important but that for reasons of data and time we were not able to examine in this paper. For example, we expect that there is a tendency to "spread solutions around" over time. Officials probably avoid an overreliance on particular solutions in successive years. Examining this and other contextual effects will be the subject of further work. It is not clear to us how general theories of budgetary behavior can ever become. It is clear that there will always be idiosyncratic elements in the contexts of budgetary behavior which will not be easily captured in theory. Our more limited aspiration is an operative simulation of problem formulation, solution and justification whose structure can be transported from city to city.

REFERENCES

Arnold, David S. (Ed.) (1981). Management Policies in Local Government Finance. Tennessee: Kingsport Press.

Bahl, Roy W., Gustely, Richard D., and Wasylenko, Michael J. (1978). "The Determinants of Local Government Police Expenditures: A Public Employment Approach." National Tax Journal 31 (1).

Bahl, Roy W., Johson, Marvin, and Wasylenko, Michael. (1980) "State and Local Government Expenditure Determinants: The Traditional View and A New Approach," in Public Employment and State and Local Government Finance. Roy Bahl, Jesse Burkhead, and Bernard Jump, Jr. (Eds.), Massachusetts: Bullinger Publishing Company.

Bailey, John J. and O'Connor, Robert J. (1975). "Operationalizing Incrementalism: Measuring the Muddles," Public Administration Review No.1, Jan./Feb.

Bolles, Albert A. (1894). The Financial History of the United States from 1789 to 1860. Reprints of Economic Classics, Volume II. New York: Augustus M. Kelley Publishers, 1969.

Booms, Bernard H. and Hu, Teh-Wei (1971). "Toward a Positive Theory of State and Local Public Expenditures: An Empirical Example." Public Finance 26 (3).

Bridenbaugh, Carl (1971). Cities in Revolt: Urban Life in America 1743-1776. New York: Oxford University Press.

Bromiley, Philip (1981). "Task Environments and Budgetary Decision Making." The Academy of Management Review Volume 6, No.2.

Burchell, Robert W. and Listokin, David (Eds.) (1981). Cities Under Stress: The Fiscal Crisis of Urban America. New Jersey: Rutgers University.

Buxton, Sidney (1888). Finance and Politics: An Historical Study 1789-1885. Reprints of Economic Classics. New York: Augustus M. Kelley Publishers, 1966.

Caiden, Naomi and Aaron Widavsky (1974). Planning and Budgeting in Poor Countries. New York: John Wiley and Sons, Inc.

Citrin, J. (1979). "Do People Want Something for Nothing: Public Opinion on Taxes and Government Spending" in D.L. Puryear et al. (Eds.) (1979) Proceedings of a Conference on Tax and Expenditure Limitations, National Tax Journal, Supplement, Volume XXXII, No.2.

Connolly, Terry (1977). "Information Processing and Decision Making in Organizations." In Staw and Salancik, (Eds.) New Directions in Organizational Behavior. Chicago: St. Clair Press.

Cowart, Andrew T., Hansen, Tore, and Brofoss, Karl-Erik (1975). "Budgetary Strategies and Success at Multiple Decision Levels in the Norwegian Urban Setting." The American Political Science Review Vol. LXIX, No.1

Crecine, J.P. (1969). Governmental Problem Solving: A Computer Simulation of Municipal Budgeting. Chicago: Rand McNally.

Cyert, Richard, James March, and William Starbuck (1961). "Two experiments on bias and conflict in organizational estimation." Management Science, 8: 254-264.

Davis, Otto A., M.A.H. Dempster and Aaron Wildavsky (1966). "A Theory of the Budgetary Process." American Political Science Review 60:529-47.

_____ (1971). "On the Process of Budgeting II: An Empirical Study of Congressional Appropriations." In R.F. Byrne et al. (Eds.), Studies in Budgeting. Amsterdam: North-Holland.

_____ (1974). "Toward A Predictive Theory of Government Expenditures: U.S. Domestic Appropriations." British Journal of Political Science 4:1-34.

Eastwood, David B. (1978). "An Adaptive Linear Expenditure System for State and Local Governments." Applied Economics 10.

Feigenbaum, Edward A. (1969). "Artificial Intelligence: Themes in the Second Decade." Information Processing 68, Amsterdam: North-Holland Publishing Co.

Feldman, Martha S. and March, James G. (1981). "Information in Organizations as Signal and Symbol." Administrative Science Quarterly 26: 171-186.

Fischer, Gregory W. and Crecine, J.P. (1980). "How Much is Enough? Explaining Presidential Defense Budget Requests." Department of Social Science Working Paper, Carnegie-Mellon University, Pittsburgh, PA.

Freedman, David, Pisani, Robert and Purves, Roger. (1978). Statistics. New York: W.W. Norton and Company.

Friedmam, Lewis (1975). Budgeting Municipal Expenditures. New York: Praeger Publishers.

Gerwin, Donald (1969). "Towards a Theory of Public Budgetary Decision Making." Administrative Science Quarterly 35.

_____ (1969). Budgeting Public Funds: The Decision Process in an Urban School District. Wisconsin: The University of Wisconsin Press.

Hayes, John R. (1981). The Complete Problem Solver. Philadelphia, PA: The Franklin Institute Press.

Hyde, Albert C. and Shafritz, Jay M. (1978). Government Budgeting: Theory, Process, and Politics. Oak Park, IL: Moore Publishing Company.

Kadane, J. and Larkey, P.D. (1980). "The Confusion of Is and Ought in Game Theoretic Contexts." Department of Social Science Working Paper, Carnegie-Mellon University, Pittsburgh, PA.

Key, V.O. Jr. (1940). "The Lack of A Budgetary Theory." American Political Science Review 34, pp.1137-1144.

Larkey, Patrick D. (1979). Evaluating Public Programs: The Impact of General Revenue Sharing on Municipal Government. Princeton: Princeton University Press.

March, James G. (1966). "The Power of Power." In D. Easton (Ed.) Varieties of Political Theory, Chapter 3.

March, James G. and Olsen, Johan P. (Eds.) (1976). Ambiguity and Choice in Organizations. Norway: Universitetsforlaget.

March, James G. (1978). "Bounded Rationality, Ambiguity, and the Engineering of Choice." The Bell Journal of Economics Vol.9, No.2.

Miller, Stephen M. and Tabb, William K. (1973). "A New Look at a Pure Theory of Local Expenditures." National Tax Journal Volume XXVI, No.1.

Mintzberg, Henry, Raisinghani, Duru and Theoret, Andre (1976). "The Structure of Unstructured, Decision Processes." Administrative Science Quarterly, Volume 21.

Natchez, Peter B. and Bupp, Irvin C. (1973). "Policy and Priority in the Budgetary Process." The American Political Science Review Volume LXVII.

Padgett, John F. (1980). "Bounded Rationality in Budgetary Research." The American Political Science Review Vol.74, No.2.

Pantaleoni, Maffeo (1883). "Contribution to the Theory of the Distribution of Public Expenditure." In Classics in the Theory of Public Finance, Musgrave, Richard A. and Peacock, Alan T. (Eds.). New York: MacMillam Company, 1958.

Patrick, Thomas M. (1980). "An Alternate Way to Look at Local Government Expenditures." Growth and Change (11:2).

Pauley, Mark V. (1975). "A Model of Local Government Expenditure and Tax Capitalization." Journal of Public Economics 6 (revised April 1976), North-Holland Publishing Company.

Peterson, Paul E. (1979). "A Unitary Model of Local Taxation and Expenditure Policies in the United States," British Journal of Political Science Volume 9, Part 1.

Pauley, Mark V. (1975). "A Unitary Model of Local Taxation and Expenditure Policies in the United States," British Journal of Political Science Vol.9, Part 1.

Pommerehne, Werner W. (1978). "Institutional Approaches to Public Expenditure: Empirical Evidence from Swiss Municipalities," Journal of Public Economics, 9 (z).

Reitman, Walter R. (1967). "Heuristic Decision Procedures, Open Constraints, and the Structure of Ill-defined Problems." In M.W. Shelley II and G. L. Bryan, Human Judgements and Optimality.

Rivera, Gerardo, Ungson, Daniel N., Braunstein, and Hall, Phillip D. (1981). "Managerial Information Processing: A Research Review." Administrative Science Quarterly Volume 26, Number 1.

Simon, Herbert A. (1977). Models of Discovery. Holland: D. Reidel Publishing Company.

Simon, Herbert A. (1979). Models of Thought. Connecticut: Yale University Press.

Sharkansky, Ira (1968). "Agency Requests, Gubernatorial Support and Budget Success in State Legislatures." The American Political Science Review Volume LXII.

Sunley, Emil M., Jr. (1971). "Some Determinants of Government Expenditures Within Metropolitan Areas." American Journal of Economics and Society 30: 4.

Turnbull, A. B. (1967). Politics in the Budgetary Process: The Case of Georgia. Ph D. diss., University of Virginia, Charlottesville.

Tversky, Amos and Kahneman, Daniel (1981). "The Framing of Decisions and the Psychology of Choice". Science, 211: pp.453-458

Wanat, John (1974). "Bases of Budgetary Incrementalism." The American Political Science Review Vol. LXVIII, No.3.

Wildavsky, Aaron (1964). The Politics of the Budgetary Process. Boston: Little Brown and Company.

Zimmerman, Dennis (1975). "On the Relationship Between Public Goods Theory and Expenditure Determinant Studies." National Tax Journal 28 (2).

Communicating with People During Emergencies

Elizabeth F. Loftus and John P. Keating

University of Washington, Seattle

Prepared for the Carnegie-Mellon Symposium on Information Processing in Organizations, October, 1981. The writing of this paper was supported in part by grants from the National Bureau of Standards, Center for Fire Research, and the National Science Foundation.

The 1980 fire at the MGM Grand Hotel in Las Vegas left 84 people dead and over 700 others injured. Within a matter of minutes, the hotel had gone from a luxurious structure to a deathtrap. In a survey conducted with survivors afterward by the National Fire Protection Association, an interesting fact emerged: most people said they wished they had been given some information on what to do. In fact, many people would not have the slightest idea what to do if they were caught in a hotel, high-rise building, or hospital when a fire, tornedo, earthquake or other emergency happened.

In any group environment, when the building codes are written, sprinkler systems installed, and the architects have finished overseeing the construction of the structure itself, the issue of safety becomes a problem of people: how to inform them of emergencies, how to evacuate them, how to forestall panic, in short, how to teach people to survive. It has been said more than once that in many emergency situations the major cause of fatalities and injuries "can be traced more to human error than to engineering failure." (Cantor, 1980, p. 3). And yet we know so little about human behavior in emergencies, except, perhaps, that people want information.

Since 1974, we have been devising ways of communicating with people during emergencies. Our work began with the development of a series of tape recorded fire instructions for the 37-floor Seattle Federal Building that eventually became a

model for federal and other high-rise buildings across the country (Keating & Loftus, 1977). Following this work in high-rise buildings, we developed a coded series of instructions for hospitals, at the request of the National Bureau of Standards (Groner, Loftus, & Keating, 1978). After this, we wrote evacuation messages for communicating with people in subways (Loftus, 1979), and have now begun work on the problems inherent in hotel environments (Keating & Loftus, 1981). Each of these environments - high-rise buildings, hospitals, subways, and hotels - posed special problems, human and technological. A number of principles began to emerge that apply to emergency communications in almost any situation.

Our first opportunity to apply psychological principles to emergency communication was provided in the Spring of 1974 when we were asked to design the evacuation system for the Seattle Federal building. In an effort to prevent a "towering inferno," this building was to be the first government building which incorporated a total safety concept, particularly when it came to fires. The basic hardware of the fire safety system was designed and several important decisions concerning the personnel evacuation were made before our involvement. While a Honeywell 316 computer controlled and monitored the bulding for many purposes, and the security staff in the basement control center scanned several television monitors, evacuation instructions for the building would be contained on cartridges that could be individually accessed. In other words, this was to be a system which included a public address component capable

of broadcasting prerecorded taped directions to communicate with building occupants who might be affected by a fire or any emergency. While vocal alarms for high rise buildings had been recommended for years by fire-safely officials, systematic development of such systems had not been attempted prior to this time.

The key to the plan was a public-address system capable of broadcasting prerecorded directions to people in different parts of the building, who would then be directed by specific vocal instruction. For guidance in the wording of such directions and their effects on behavior, we reviewed basic research in social, cognitive and human-engineering psychology. Few studies specifically addressed evacuation messages specifically. But with this literature as a guide, we developed several different kinds of instructions.

When a fire is reported, the computer determines the location of the fire floor via smoke detectors or call in boxes. The elevators are immediately captured and sent to the ground floor. A message from a cartridge is played through the speakers in all 22 elevators, explaining that there is a fire (or other emergency) that the elevators are returning to the lobby floor, and that passengers should disembark and await instructions. Immediately after this message, several messages are transmitted to people on the affected floors.

Suppose a fire breaks out on the 20th floor of a building. People on floors 19 and 20 would be asked to go down to floor 18, while those on floor 21 - one floor above the fire - would

be asked to go up to floor 22. People on any floors above 21 need not be evacuated in most cases. For the fire floor, we designed this message:

Female voice: "May I have your attention please. May I have your attention please."

Male voice: "There has been a fire reported on the 20th floor. While this report is being verified, the building manager would like you to proceed to the stairways and walk down to the 18th floor. Wait on the 18th floor for further instructions. Please do not use the elevators as they may been needed. Please do not use the elevators, but proceed to the stairways."

The message tells people exactly what is happening and what they are supposed to do. Several aspects of the message should be noted. First, all essential information is repeated. People hear twice that they should go to the stairways, that the 18th floor is their destination, and that they should not use the elevators. Various research studies have shown that repetition of a message promotes understanding and retention (e.g., Hebb, 1961; Hellyer, 1962). Furthermore, the message contains relatively common words. Numerous empirical studies have shown that words that are used commonly in the language are also more easily understood (Howes, 1957). This point has been included in one authoritative source, the Human Engineering Guide to Equipment Design (Van Cott & Kinkade, 1972), which stated "other things being equal, the more frequently a word occurs in everyday usage, the more readily it is correctly identified when transmitted over a speech communication system." (p. 219).

Occasionally a less common word, like "proceed", is used in favor of a more common word, like "go", when the redundancy in the multi-syllable words would aid comprehension.

The statement "The building manager would like you to proceed...." gives the impression that someone is in control of the situation. When an emergency arises, people tend to follow the lead of a person who shows initiative or who appears to be in control (Phillips, 1975). If the "advise" is bad, problems can arise, and was evident in the classic example during the evacuation of Nagasaki. Hundreds of survivors attempted to leave the area along a rough footpath, beset with rocks, roots, and branches that snagged the clothing and tore at the skin. They toiled along this path up steep inclines and through a good deal of mud, continuing to follow a man who led them. Next to this path was a smooth, level road that would have been far easier to use. When questionned as to why they used the inferior path, not one person had a good reason. Rather, they said that they presumed that the man who was leading them knew what he was doing. Although orders from above must be given cautiously or they will not be heeded (Quarantelli & Dynes, 1972), if done properly they can be enormously effective. Another aspect of the high-rise message is that it specifies that a fire has been reported, avoiding the more ambiguous term "emergency". The use of the less ambiguous terminology was suggested by research in social psychlogy which examined how people respond in simulated emergencies (Latane & Darley, 1970). It was the conclusion of this line of research that when people

are in groups they tend to shunt responsibility for action to other members of the group, a response that especially occurs during ambiguous situations. It has been estimated that approximately 20% of the population may perceive an ambiguous emergency situation improperly and resort to counterproductive behavior (Phillips, 1973). Informing people of the true situation - fire, in this case - further reduces the chance that people will create vivid rumors, increasing chances for counter-productive behavior.

One important aspect of the message is that people are given a reason for engaging in particular behaviors. They are told, specifically, not to use the elevators because "they may be needed." In many areas of research it has been shown that people are more likely to follow suggestions if they are given a good reason to do so. This was found, for example, in a study of control of littering behavior (Geller, Whitmer & Orebaugh, 1974). These investigators showed that general instructions ("Please dispose of litter properly") reduced the number of handbills left on shelves, counters and floors more than 50%. More to the point, specific instructions ("Please deposit in green trash can") were no more effective, but adding a reason for depositing the handbill in the trash can ("Please dispose for recycling in green trash can") made the most effective appeal. Similar beneficial effects of supplying a reason for particular behavior have also been found by Latane and Darley (1970).

Finally, note that our high-rise message is introduced by a

female voice, and the instructions themselves are delivered by a male voice. Research suggests that switching from a female to a male voice will be noticed even when people are not really paying attention (Cherry, 1953; Cherry & Taylor, 1954). One might reasonably ask why a male voice did not introduce a female, since the basic work of Cherry and others has shown that the switch is noticed in both directions. The reason for choosing the male voice to deliver the bulk of the directions is that in the early 1970's in American society, males were stereotypically looked to as the ones who take charge in an emergency. Relying on this stereotype, however, unjustified it may be, led to the recommendation of a male voice for most of the directional delivery. The choice may be superior for other reasons, as well, however. Recent research has shown that people comprehend and remember more of a message when it is delivered by a male rather than female voice (Gruber & Gaebelein, 1980). Furthermore, the decision is wise from the point of view of the elderly. With advancing age, certain structural changes occur that affect both visual and auditory processing (Fozard, Wolf, Bell, McFarland & Podolsky, 1977; Hoyer & Plude, 1980). For example, the diameter of the pupil decreases, affecting vision, and the sensitivity to various wavelengths, particular high frequency wavelengths, is reduced, affecting audition. Since females tend to have higher frequency voices, they will become relatively more difficult for older people to understand.

One challenging aspect of moving people in the Seattle

building was that people working on the floor above a fire would be required to go up a flight of stairs. Most people, by natural inclination and a lifetime of conditioning, want to go down and leave a burning building. Fire officials, however, want to prevent traffic jams in stairwells leading from the fire floor, where quick exit is urgent. We worried that people asked to go upstairs against their best instincts might not follow this instruction. Consequently, we added a phrase to the standard message, delivered only to those who are asked to go up, that the floor above them is safe.

To summarize, we have briefly indicated some of the decisions we made, and the rationale for those decisions. Many other considerations influenced our decisions, although we have not had the time to enumerate these here. In brief, they involve attention to semantic and syntactic features of Engish that are known to impede comprehension. In the semantic domain, we tried to minimize the use of technical terms, polysyllabic words, unusual prepositional phrases, excessive formality, and vagueness. In the syntactic domain, we tried to minimize the use of nominalizations, passive constructions, conditional constructions, multiple negatives, embedding of many subordinate clauses. These considerations have previously been helpful in the constuction of comprehensible jury instructions (Charrow & Charrow, 1979).

Since in many cases the leap from social science research to the emergency situation was rather large, it seemed imperative to test the plan. Accordingly, our first test was

conducted shortly after the occupation of the new building. To lend realism to the test, firefighters in full regalia were stationed conspiciously on all floors being evacuated. Drills were conducted in two sections of the building, and when completed, participants filled out questionnaires designed to measure their reactions to the system. The results were encouraging: All floors targeted for evacuation were vacated within one and a half minutes of the start of the drill. Behavior was so orderly that a chief of the Seattle Fire Department was prompted to say it was the smoothest fire drill he had witnessed in his many years with the department.

Since 1974 several new federal buildings have adopted variations of the Seattle messages. It thus became possible to compare the voice systems with the traditional bell alarm still used in older federal buildings. We selected several federal office buildings around the country, some with voice messages (in Chicago and Roanoke, Virginia) and others with bell systems (e.g., Cleveland, Kansas City). We held fire drills in each building, and compared data from interviews with participants. Unfortunately numerous unanticipated problems arose, precluding the detailed comparisons that we had hoped for. For example, in some buildings, sensitive government agencies (FBI, IRS, CIA) were apprehensive about evacuating their offices. In a few of the buildings with prerecorded messages, the equipment malfunctioned. Despite these problems, some tentative results emerged. In buildings devoid of technical problems, people had no trouble understanding the instructions, and virtually all

thought the reason for the evacuation was clearly stated. Although in most buildings, evacuation was complete within five minutes, in buildings with well-functioning prerecorded voice systems, participants completed their evacuation with two or three minutes.

In sum, we have described a case study in which psychological principles were applied to the real life problem of communicating with people during emergencies. Although the success of this application has been confirmed only in fire-drill type situations, the outcome of these tests are encouraging.

We now turn to a different environment altogether - a metropolitan subway. When the Bay Area Rapid Transit System (BART) asked us whether we could develop standard messages for subway evacuation, we wondered whether the ones we had written for high-rises would be applicable. This opportunity came about in early 1979 when a fireman was killed aboard a BART subway train in the tunnel under San Francisco Bay. The fire made officials realize how little attention had been devoted to making plans for orderly evcuations of the subway during emergencies. The California Public Utilities Commission ordered the tunnel closed until BART authorities could demonstrate that the subway system was safe. Almost immediately, one of the city's leading engineering firms was asked to prepare a comprehensive safety plan, and as part of this plan it sought the help of a specialist in information processing who could develop a set of instructions to be given by the trainman over

the intercom in case of future disasters. Based upon our experience with the high-rise buildings, we developed a message that was significantly shorter and clearer than previous versions that BART had experimented with. Our message began this way:

"May I have your attention please. A fire has been reported on this train. BART Central would like you to leave the train for your own safety. Please leave the train now and walk slowly to the opposite tunnel. Follow these instructions for going to the opposite tunnel...."

Each word and phrase in the message had a purpose. As in the case of the high-rise messages, at the outset, this message tells passengers exactly what is happening: "A fire has been reported on this train." The term BART Central" gives them the feeling that there is an authority in control who will tell them what they are to do and why. The next line, "Please leave the train now and walk slowly to the opposite tunnel" repeats the instruction to leave the train....and so on. With the new streamlined instructions, the evacuation plan was deemed acceptable by the Commission, and service was soon resumed on BART.

Could these basic messages be used in any environmental setting? When the National Bureau of Standards asked us whether we could develop standard messages for hospitals, we wondered whether the ones we had written for the high-rises or for the subway would be applicable. The answer was no. Our earlier messages had been based on the important principle of

reducing ambiguity. If a fire had broken out, people were told it was a fire. However, the hospital posed a unique evacuation problem. When a fire occurs in a hospital, information is typically communicated to doctors, nurses, and other personnel, but withheld from patients. The response of recent burn victims, patients and visitors in intensive-care units, and the mentally ill is especially unpredictable in such emergencies. Thus, most hospitals use messages that are "disguised" or coded to be purposely ambiguous.

The major problem with coded messages in actual use is that the particular code used by any given hospital seems to be almost randomly chosen. In some cases "code red, four west" alerted appropriate personnel to a fire on the fourth floor of the west wing. Other hospitals might use "code blue" or even "smoky the bear". Other color and number codes might indicate a cardiac arrest, a bomb threat, or some other emergency altogether. Key phrases in these messages often bore no relation whatsoever to the emergencies they announced, and they differed from institution to institution. Thus doctors, nuerses and other staff were often confused. In one fire drill that we observed at a major metropolitan hospital, a doctor turned to a nurse and asked: "Is code red for fire or cardiac arrest?"

To reduce the confusion, we developed a standardized message that would be easily grasped (Groner, Loftus, & Keating, 1978). We needed some language that could be unambiguously associated with fire, but which we could then disguise by preceding it with the title "nurse" and announcing it as a

surname. Using standard free association procedures, we determined that the word "blaze" was ideal; it elicited "fire" as a primary response in most respondents, and it was considered to be a reasonably plausible yet rare surname. We recommended a disguised message such as "Nurse Blaze, four west" be used to communicate to targeted personnel that a fire had been reported on four west. Aside from its other useful attributes, it was rather unlikely to lead a hearer to ask: "is Nurse Blaze for fire or cardiac arrest?" This suggestion is under serious consideration by the National Bureau of Standard's Life/Safety Fire Program.

We have begun to turn how to the problem of communicating with people in hotel emergencies. As in other situations, a major problem is that people do not know what to do when a fire or other emergency occurs. For example, it does not occur to many people that they should take their key with them when they leave their room during a hotel fire. Many people do not know that the risks of jumping out of a third story window are usually greater than staying put in the hotel room and fighting the smoke. Many people do not under stand how dangerous it is to use the elevators in a fire (Keating & Loftus, 1981).

We have tried to apply the principles used in devising other messages to the hotel environment. Again, a number of principles apply across situations, but the hotel situation, too, poses some unique problems. Hotel guests are generally less familiar with the overall design of a hotel where they may be staying for the first time than the overall design of an

office building where they may have been working for years. The oblem of an individual attempting to use the elevators, the "dominant route" (Keating, 1981), the only route that has ever been taken, may be far greater in the hotel than the high-rise office building. Naturally, special problems will require special solutions.

As we move to disasters other than fires, the psycholocal literature offers even less guidance about the proper wording of evacuation instructions. Warning people about bomb threats, for ample, presents its own anomalous problems. In the Seattle Federal building, if the emergency is a bomb threat, personnel in the building are directed to proceed to the nearby ferry terminal building and wait for further instructions, or leave the building and go home for the day. The entire building is affected, rather than simply portions of the building, and these differences need to be taken into account. However, some inciples would be the same: any message shuld be short, clear, use familiar language, repeat important information, cite authorities are are in control.

When more research is done to discover what counterproductive behavior people ordinarily engage in, better evacuation messages can be designed. For example, in earthquakes, many people inside a building promptly run outside. This is the worst thing they can do, according to many safety perts, since dashing for an exit usually increases the risk of being hit by falling bricks, plaster, and shattered glass. Rather, people should get under a strong desk or table, or in an

interior doorway. Prerecorded messages in public buildings might profitably be used to counteract the instinct to run outside, and inform them as to proper behavior. One problem that is unique to the earthquake emergency is that often there is not much advance warning. Officials may learn of an impending quake only 15 seconds or so before it occurs, and thus it may not be possible to deliver a prerecorded message prior to the onset of the quake. However, assuming the damage were not especially severe, such messages could be delivered afterward to instruct people on the proper behavior during the aftermath.

In an emergency-filled world, rigorous studies of how to communicate with people could be one of the most useful fields for applied psychology. The present research is simply a beginning. It has generally involved a leap of faith in applying basic principles to an real-world problem, and in only some of the cases have tests of these applications been possible. Generalization to a new situation can often be risky, and this is something we recognize. Although many papers end with a plea for more research, it seems particularly appropriate to do so in this case. Clearly much more work is needed before we can fully understand how to communicate with people in difficult situations.

References

Cantor, D. (Ed.) Fires and Human Behavior. New York: Wiley, 1980.

Charrow, R. & Charrow, V. Making legal language understandable: A psycholinguistic study of jury instructions. Columbia Law Review, 1979, 7, 1306-1374.

Cherry, E.C. Some experiments on the recognition of speech with one and two ears. Journal of Acoustical Society of America, 1953, 25, 975-979.

Cherry, E.C. & Taylor, W.K. Some further experiments on the recognition of speech with one and two ears. Journal of Acoustical Society of America, 1954, 26, 554-559.

Fozard, J.L., Wolf, E., Bell, B., McFarland, R.A., Podolsky, S. Visual perception and communication. In J.E. Birren & K.W. Schaie (Eds.) Handbook of the psychology of aging. New York: Van Nostrand Reinhold, 1977.

Groner, N.E., Loftus, E.F., & Keating, J.P. Calling nurse blaze: Tailoring programs to fit human behavior. Hospitals: Journal of the American Hospital Association, 1978, 52, 111-115.

Geller, E.S., Whitmer, J. & Orebaugh, A. The power of print. Behavior Today, 1974, 5, 245.

Gruber, K. & Gaebelein, J. When a man talks, you listen. Psychology Today, Dec. 1980, 25-27 (based on work appearing in Sex Roles, Vol. 5.)

Hebb, D.O. Distinctive features of learning in the higher animal. In J.F. Delafresnay (Ed.) Brain mechanisms and learning. New York: Oxford University Press, 1961.

Hellyer, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. Journal of Experimental Psychology, 1962, 64, 650.

Hoyer, W.J. & Plude, D.J. Attentional and perceptual processes in the study of cognitive aging. In Poon, L.W. Aging in the 1980's. Washington, D.C.: American Psychological Assn, 1980.

Keating, J.P. The myth of panic. Paper presented at the U.S. Fire Administration Conference, San Francisco, 1981.

Keating, J.P. & Loftus, E.F. Vocal alarm systems for high-rise buildings - A case study. Mass Emergencies, 1977, 2, 25-34.

Keating, J.P. & Loftus, E.F. The logic of fire escape. Psychology Today, 1981, 15, 14-19.

Latane, B. & Darley, J. The unresponsive bystanderj: why doesn't he help? New York: Appleton-Century Crofts, 1970.

Loftus, E.F. Words that could save your life. Psychology Today, 1979, November, 102-110, 136-137.

Phillips, A.W. People, physiology and fire problems. Paper presented at the NSF/RANN Conference on Fire Research, Harvard University, 1975.

Phillips, A.W. The physiological and psychological effects of fires in high-rise buildings. Factory Mutual Record, May-June, 1973, 8-10.

Quarantelli, E.L. ' Dynes, R.R. When disaster strikes. Psychology, Today, 1972, February, 67-69.

Van Cott, H. P. & Kinkade, R.G. (Eds.) Human Enineering Guide To Equipment Design. Washington, D.C.: American Institutes for RESEARCH, 1972.

GOSSIP, INFORMATION, AND DECISION MAKING

by

James G. March and Guje Sevón

Graduate School of Business
Stanford University

Swedish School of Economics
and Business Administration
(Helsinki)

To appear in: Lee S. Sproull and Patrick D. Larkey, eds., Advances in Information Processing in Organizations, Vol. I (Greenwich, CT: JAI Press), forthcoming, 1982.

This paper was prepared for a symposium on information in organizations held at Carnegie-Mellon University, October 16-17, 1981. We are grateful for the comments of Nelly de Camargo, Omar El Sawy, and Johan Olsen.

ABSTRACT

This paper examines the relation between information and decision making, particularly in organizations. We observe that much of the information that human beings seek and receive is gossip, that is, information without decision relevance. We ask two general questions: Why do we observe so much idle talk in life? And what are the implications for understanding organizational decision making and the design of management information systems? We conclude that the prevalence of idle talk stems from some systematic ways in which ordinary life, including ordinary managerial life, differs from the life anticipated by a focus on decision making, and that information engineering may, as a result, be somewhat less informed by decision theory than we sometimes expect and somewhat more informed by literary criticism and the philosophy of education.

1. THE VALUE OF INFORMATION AND INVESTMENTS IN GOSSIP

In discussions of the design of information systems in organizations, the value of information is ordinarily linked to management in a simple way. We imagine that management is primarily a matter of making decisions, and that a decision maker chooses among several alternatives on the basis of information about consequences and preferences that are conditional on a choice. Additional information has value to the extent to which it can be expected to affect the choice. Thus, a prediction of snow in Helsinki has no value to a road maintenance crew in New York, as does a prediction of snow in New York in July. In the first case, the information is irrelevant to any decision; in the second it is redundant with prior information.

Within a decision theory frame, investments in information sources are made up to the point at which the marginal expected cost of the source equals the marginal expected improvement in deci-

sions; and information systems are designed to assure that scarce resources of money and attention are allocated efficiently from such a point of view. The value of information depends on the decisions to be made, the precision and reliability of the information, and the availability of alternative sources (Marschak and Radner, 1972; Hirschleifer and Riley, 1979). Although calculating the relevant expected costs and returns is rarely trivial, the framework suggests some very useful rules of thumb: Don't pay for information that cannot affect choices you are making. Don't pay for information if the same information will be freely available anyway before you have to make a decision for which it is relevant. Don't pay for information that confirms something you already know. In general, we are led to an entirely plausible stress on the proposition that allocation of resources to information gathering or to information systems should depend on a clear idea of how potential information might affect decisions. Who needs the information and how is it relevant?

Actual investments in information and information sources appear to deviate considerably from these conventional canons of information management. Consider, for example, the daily newspaper. Significant numbers of individuals and institutions, including many located considerable distance from New York, purchase and read the New York Times. The newspaper provides some information that is potentially relevant to some decisions faced by its readers: sche-

dules of events, advertisements, reviews of books and various performances, market information, betting odds, etc. But what is the likely decision relevance of the information on sporting events? Or of the news from Washington, London, and Tokyo? Much of the political news we devour, like much of the news from Hollywood, is essentially gossip. It may have relevance to some decisions made by major political actors, but it is hard to see any analysis of the decision situations of professors in Palo Alto that would lead to an investment in the political or sports news of the New York Times, or indeed to much of the other content if the primary reason for such an investment were information value in standard decision terms (Simon, 1967).

The daily newspaper is only a mundane illustration of a more general phenomenon. Even in job situations with a good deal of task specificity, we devote substantial time to gathering and transmitting information that has no obvious connection to the immediate decisions we contemplate. Individual and organizational consumers of information invest in decision support systems and patronize reports and conversations with little apparent attention to their decision relevance. Studies of the use of research in the public sector identify little connection between the research information that agencies seek and the decisions they make (Rich, 1977; Weiss, 1977; Deshpande, 1981). Business firms, armies, hospitals, and other organizations we observe

systematically gather, store, and display information that they do not, indeed could not, use; they invest in large information systems and in irrelevant forecasts (Swanson, 1978; Feldman and March, 1981). Although there is no question that individuals and organizations invest in decision-relevant information under many circumstances, much of the information that is gathered and reported makes little direct contribution to resolving choices. In that sense at least, it is essentially idle talk.

2. UNDERSTANDING IDLE TALK

The persistence and pervasiveness of idle talk makes it relevant not only to understanding everyday life but also to improving managerial behavior and to designing management information systems. Gossip cannot easily be ignored. It is either an inefficiency in information systems or a symptom of inadequacy in our ways of thinking about information. Without denying the former possibility, we wish to explore the latter. The information investments of individuals and organizations seem to suggest that the connection between information

and action is more subtle than we have made it, or that our focus on choice as the central metaphor of life is misleading. We want to argue that both things are true and that they have implications for understanding and improving the uses of information in organizations.

Research on gossip is primarily concerned with information, with or without a known basis in fact, about the personal affairs of individuals (Hannertz, 1976; Rosnow and Fine, 1976). Gossip is sometimes seen as a simple source for entertainment (Rosnow and Fine, 1976). However, much of the research portrays gossip as contributing to system maintenance more than to decision making. It is seen as a way to communicate rules, values, and morals -- usually by pointing at failures to satisfy them. It facilitates the diffusion of community traditions and history (Lumley, 1925; Gluckman, 1968; Haviland, 1977), and the maintenance of exclusivity (Gluckman, 1963). Gossip is a way of making friends (Rosnow and Fine, 1976), a way of protecting personal interests (Paine, 1967, 1970), and a way of legitimizing collective action, such as a riot (Mitchell, 1956). Though it often reinforces existing beliefs by providing an interpretation of ambiguous experience that is consistent with them (Allport and Postman, 1947), and offering a guide to existing social structure (Hannertz, 1967), gossip is also a vehicle for social change. It is a mechanism for a collective reconstruction of reality in which existing explanations of the nature of things are modified and new sensibilities and ideas emerge and are ela-

borated (Shibutani, 1966).

It is clear that such research on gossip can hope to provide only indirect clues to the analysis of idle talk in organizations, but it suggests that there may be somewhat more intelligence in the social processing of decision-irrelevant information than a decision theoretic analysis would indicate. Although our search for reasons for idle talk will extend beyond the relatively narrow focus of gossip research, it is in the same general spirit. We will argue that an exclusive focus on the role of information in well-defined decisions is likely to lead to an inadequate characterization of the information investment problem and thus to inadequacies in the design of management information systems. In particular, we will argue that:

(1) Information systems need to be exercised in slack times in order to be useful when needed. Information is processed, in part, to maintain the system rather than to use it.

(2) Human action is often less a matter of choice than a matter of imitating the actions of others, learning from experience, and matching rules and situations on the basis of appropriateness.

(3) Decisions are often made in situations that are quite distant from the situations implicit in ideas of rational choice. Neither the precise decisions, the alternatives, the objectives, nor the causal structures are clear.

(4) Information is often as much dedicated to developing interpretations, explanations, understandings, and enjoyments of the events of life as it is to resolving specific choices.

Gossip as system maintenance. Idle talk is a way an information system is kept effective. On the one hand, it smooths interpersonal and inter-group strains introduced by organizational life. People engage in idle talk in order to exhibit their reasonability and legitimacy, to exchange sentiments of solidarity, to reduce the risks of misunderstanding, and to make it easier to arrange the minor flexibilities that allow an organization to function (Frankenberg, 1957; Gluckman, 1963; Rosnow and Fine, 1976). Arguments that Finnish business organizations need the *sauna* in order to thrive, or that American businesses need the three-martini lunch, may be as fatuous as they appear; but coordination in families, neighborhoods, societies, and organizations is facilitated by the gossip that fills such institutions. This integrating and catalytic consequence of talk is hard to link concretely to specific actions or specific consequences, but it

appears to almost all observers of social systems to be relatively fundamental.

Moreover, a communication system may need irrelevant exercise to maintain effectiveness. individuals, organizations, and species risk developing specializations that reflect optimal short-run allocations of effort and ignore long-run investments in capabilities for dealing with infrequent or unlikely situations of importance. From this perspective, the specific content of talk is largely irrelevant. Gossip maintains links among people, for example in a neighborhood or between organizations, during those long periods when communication is unneeded, so that the communication links will be easily available should they be needed. Similarly, an organization may find that idle talk has the consequence of maintaining connections among parts of the organization that require few regular connections. Idle talk of this sort is an inexpensive substitute for emergency drills. From such a point of view, the justification for idle talk comes not from some subtle relevance but directly from its irrelevance. If it is desirable to maintain links among parts of an organization that are, under normal conditions, quite sensibly not connected, the decision irrelevance of gossip has the admirable property of producing contact between parts of an organization having no current need for coordination.

Information and alternative concepts of action. The idea of choice and the idea of expected value maximization with which it is joined in contemporary decision theory are possible metaphors for thinking about action in organizations; but they are not the only possible metaphors. The idea of decision making implies an anticipatory, consequential logic. That is, it assumes that action results from two guesses about the future: a guess about the uncertain future consequences of taking one action or another, and a guess about the uncertain future feelings a decision maker will have about those consequences when they are realized. Such a vision seems often to be a useful one for understanding some parts of organizational action, but students of organizations generally note several other ways in which organizational actions might be interpreted (March, 1981).

It is possible to see action as reflecting experiential learning in which propensities for doing one thing or another change as a result of simple behavioral reinforcement. In such a view, action is history dependent rather than expectational; the relevant information is information on contemporaneous events or past experiences rather than forecasts of the future. It is possible to see action as reflecting contagion, as spreading through a population of actors like measles through a population of children. In such a view, action diffuses on the basis of exposure and susceptibility; the relevant information is information on what other people are doing. It is possible to see

action as rule following, as the matching of rules, procedures, and routines to appropriate situations. The routines may be seen as having evolved, or been learned. Experiential history is stored in the rules and cannot easily be retrieved in uncoded form. As a consequence, the logic of routine action is classificatory rather than consequential or anticipatory. It is filled with calculations of appropriateness, and the relevant information is information that maps a set of rules for action onto a situation.

These alternative metaphors for action have been used extensively in interpreting individual and organizational behavior. They imply a different conception of relevant information than that based on a conception of anticipatory choice. In general, the information requirements for learning, diffusion, and rule, procedure, and routine following place a greater emphasis on knowing what has been happening in the past, or is happening now, than do the requirements for decision making, and a lesser emphasis on forecasting the future. The cognitive questions they ask involve description and classification more than they do chains of conditional consequences. The dominant vision is one that sees an organization as monitoring the environment for surprises (threats or opportunities) rather than assessing alternatives. Some information that looks like idle talk within the frame of intentional, anticipatory choice is more relevant to action when it is seen within these other metaphors.

The decision context of information. Some decisions in organizations are readily amenable to a tight linkage between flows of specific information and the making of specific decisions. The chatter between an aircraft pilot and a air traffic controller, for example, normally contains very little idle talk. Information is precoded in decision-relevant form. Similar situations are common throughout modern organizations, as are successful efforts to design and maintain sensible information inventories. An optimal information inventory can be determined as long as it is possible to make reliable and precise predictions of future decision deadlines and information requirements. In a stable, uncomplicated environment, such forecasts will often be accurate enough to connect the collection of specific information to specific future decisions. We would expect an optimal decision support system to develop an inventory of information that has, relative to its cost, a reasonable chance of being useful in future decisions.

There are, however, other kinds of decision contexts; and it seems at least possible that the pervasiveness of gossip is due, in part, to the ways some decision contexts lead to loose linkages between specific current decisions and specific current information. First, there is no necessary reason to expect that decision deadlines will be consistent with the timing of information. One reason for an organiza-

tion to gather information that is irrelevant to immediate choices is the disparity between the time (or other resources) required to obtain the information and the time that will be available when a decision using the information has to be made. If it takes a relatively long time to assemble and interpret information, it may be necessary to invest in information inventories in anticipation of the future stream of decisions. Decisions in modern warfare are an obvious example of a situation in which the real time demand for decisions in battle may easily overload an amry's information gathering capabilities unless substantial information inventories are developed in advance. Planning for emergencies in general involves building inventories of information and routines whose relevance depends on a possibly unlikely contingency. Similarly, the work of collecting and organizing information about customers, competitors, friends, and enemies often cannot reasonably be postponed until the specific decisions involved are immediate.

Under such circumstances, it is possible that the "irrelevance" of many investments in information is a post hoc illusion. Normally, we commit ourselves to attending to a source with only an estimate of what will be said; we invest in a spy system with only an estimate of what it will produce; we buy an econometric forecasting service with only an estimate of what information it will generate. If only a few possible signals are important to decision, but they are very important and very unlikely, most of the information actually received from

a decision relevant source will appear to be decision irrelevant.

Second, where the future stream of decisions is unclear and future preferences are ambiguous, the selection of an information inventory is likely to be difficult. Returns in terms of improved future decisions are hard to assess when the long-run decision stream is not well-specified. We do not know what we might need to know. This ambiguity tends to make the collection of information more dependent on properties of the information available than on predictions of possible future decision contexts, and leads to the accumulation of knowledge of of unknown decision relevance. Since estimates of the costs of gathering and storing information are less affected by decision uncertainties than are estimates of the benefits, calculations of the net return from alternative information strategies are likely to be primarily sensitive to variations in costs. For example, organizations are likely to gather considerable information of dubious benefit to decision makers when it is possible to transfer the costs of gathering it to another budget, as in the case of government agencies that require others to collect data for them, or the case of central office functionaries in dealing with district offices. If information inventory decisions are primarily a function of the costs of information gathering and processing, idle talk is likely to be an attractive information system for many organizations and many individuals. It provides information in a timely and inexpensive way. The information may or may not turn out to

be relevant in the long run, but future decision needs are sufficiently unclear as to make such uncertainties characteristic of almost any information that is available.

Third, a loose link between decisions and information can be strategic. It is common in talking about the design of information systems to disregard conflicts of interest between information sources and decision makers, or to assume such conflicts are managed through explicit principal-agent contracts that assure jointly consistent behavior. In this way, strategic manipulation of information can be ignored. Where such contracts are difficult to specify or conclude, however, innocent information sources cannot be assumed. A request by a decision maker for information is a signal of decision relevance and thereby an invitation for information sources to try to manipulate the content or increase the implicit price of information. Consequently, it may be useful for a decision maker to obscure information relevance, to encourage the free flow of mostly irrelevant information in order to reduce the precision with which decision consequences can be anticipated by information sources. Such a strategy does not reduce the incentive toward lying by information sources, but it limits the potential for lying effectively and increases the innocence of the unwittingly relevant information that is provided.

Information and interpretation. At most, the New York Times is elevant to only a few decisions most of us can expect to make in the foreseeable future. It is filled with gossip about politics, sports, art, and finance that is distant from choices that we face. Yet, we think it possible that the New York Times is, nonetheless, useful to our lives. It provides the story line for our pretenses and the content of our conversations. To see life, or management, as decision making is to see it inadequately; and one of the reasons that much of the information that is communicated in organizations, as in life, is not obviously relevant to decision making is that choice is not as compelling a metaphor for managers or other individuals as it is for students of choice. From many points of view, individual and organizational life is better seen as dedicated to developing interpretations of events and understandings of history than as making choices.

Intelligent choice often presumes understanding, of course, and it is possible to see the interpretation of history as instrumental to the action (choices) by which we seek to control our fate. It can, however, be seen as more fundamental than that. Perhaps, interpretation is more a primary feature of human behavior than a servant of choice. From such a perspective, information is sought and considered because it contributes to understanding what is going on in life; and understanding what is going on is important independent of any purpose to which the knowledge might be put. Perhaps, we can better understand

the uses of information in organizations if we see information, and decision making, as part of an effort to comprehend and appreciate human existence, as driven by elementary curiosity as much as a hope for instrumental advantage. Information is not gathered in preparation for life. Gathering information is life. Moreover, this process of appreciation, of discovering, elaborating, and communicating interpretations of events, is a pleasure. Individuals and organizations are entertained by exchanging information and constructing what might be imagined to be true, or just imagined. Fantasy is a part of understanding, and idle talk is fun.

Such a viewpoint is not novel. It is familiar to literature, as well as anthropology and sociology, less familiar to economics. We can see organizations as having been designed (or evolved) around some problems of developing, enjoying, and sharing interpretions of reality, communication in organizations as tied to the discovery, clarification, and elaboration of meaning, and the process of decision making in organizations as a performance within which individuals and groups construct an interpretation of experience that can be shared meaningfully and enjoyably.

Because understanding what is going on is important, people who understand what is going on are viewed as people of importance. People who know what is going on are eager to exhibit that fact, and

the exchange of information is the exchange of signals about power, position, and competence. Note that from this perspective what makes information a source of power is not any added capabilities for effective action that knowledge provides, but rather the simple possession of a scarce and valued resource and the capability for signaling individual and organizational significance. Information is exchanged for other information (or other goods); it is exhibited as testimony to worth, much the way a plutocrat exhibits wealth.

If we see decisions as being somewhat less central to life than they are to decision theory, it should not be surprising that a theory of information <u>for decision making</u> finds parts of information life irrelevant. And if the purpose of a formal information system is to strengthen the information base of current decision making, it may not be necessary to be overly concerned with these "irrelevent" considerations. If management can be seen as decision making in a decision theory mode, then there is no substantial difference between the idea of a good management information system and the idea of a good decision support system. Designing the former consists in designing the latter.

However, if we take a more general perspective on management, the design implications of these other factors may not be trivial. If we relax the presumption that the primary interpretive theme of

management is choice and thus that the quality of management is determined primarily by the intelligence of the managerial decision making, we may conclude that we have been excluding some important things in thinking about information in organizations. If management is seen as involving discovering new objectives, developing myths and interpretations of life, and modifying the diffuse beliefs and cultural understandings that make organizational events comprehensible, and life enjoyable (March, 1973, 1978, 1980; Jonsson and Lundin, 1977; Feldman and March, 1981), then it is not obvious that the best management information system is a decision support system. Intelligent managers might pay more for, and attend more to, something different.

3. MANAGEMENT INFORMATION SYSTEMS AND EDUCATION

If the arguments made in the preceding section have merit, then it is not hard to see why the gathering and processing of information in organizations would involve a significant amount of idle talk. At the same time, there seems little reason to assume that the chorus of idle talk that we hear in organizations is optimal. We require some

way of approaching the design of information systems that is sensitive to the sensibility of idle talk, yet seeks to improve the quality of information available. We have tried to argue that a decision-oriented view of information, however valuable within the context of decision making, may sometimes be misleading as a more general base for understanding and improving a management information system. If the innocence of information cannot be assumed, or if the future stream of decisions cannot be anticipated well, or if action is based on matching behavioral rules to appropriate situations, or if the point to information is the elaboration of a system of meaning, organizational engineers need a somewhat different set of models for designing information systems.

As always, it is easier to see the limitations of a decision perspective on information than to develop a clear alternative. One major problem, of course, is that the several complications we have identified do not immediately suggest a common remedy. If we focus on the complexities of the information context of decision making while still retaining a general decision theory frame, we can generate a set of suggestions that emphasize exploiting reductions in the costs (or improvements in the speed) of gathering, storing, and retrieving information, working on reducing response time and retrieval time rather than anticipating specific needs. If we focus on alternative ideas for examining action, we generate a set of suggestions that em-

phasize monitoring the environment for critical signals and surprises, working on our capabilities for timely notice of environmental events rather than analytical or expectational capabilities. If we focus on the uses of information in the construction of meaning, we generate a set of suggestions that emphasize the ways in which meaningful stories are constructed and shared and the understanding of experience, working on the flexibility and imagination of the system for creating and articulating interpretations.

These different ideas lead to different implications, not obviously mutually consistent and not trivial to accomplish. It would appear that we require some notion of the value of alternative information sources that is less tied to a prior specification of a decision (or class of decisions) than to a wide spectrum of possible decisions impossible to anticipate in the absence of the information; less likely to show the consequences of known alternatives for existing goals than to suggest new alternatives and new objectives; less likely to test old ideas than to provoke new ones; less pointed toward anticipating uncertain futures than toward interpreting ambiguous pasts. The requirements considerably exceed our capabilities. We neither understand idle talk well enough nor are rash enough to propose a precise alternative model for the design of information systems in organizations. Our objective is more timid, to propose some caution in treating the problems of organizational information as problems in improving decision

making, and to suggest that alternative perspectives are not completely ..lien to our intellectual traditions.

In fact, a view of information and life not far from the one we have sketched is a quite traditional one, associated classically with literature, art and education; and if there are appropriate models for a management information system of this sort, perhaps they lie in discussions of the nature and design of education rather than in modern theories of decision. Perhaps management information designers could profit from some attention to the ancient and modern discussions of the linkage between education and life, the arguments over the relevance of "relevance" in thinking about a curriculum, and the efforts of art and literary criticism to explicate the expression of meaning.

To be sure, there are differences between an organization and a society and between managers and educators or artists. Many of those differences involve the relative specificity of activities and objectives in organizations, compared with the relative diffuseness of broader social relations. The differences make the leap from the analysis of education to the analysis of organizational information a large and possibly treacherous one. But not entirely foolish. As we discover the elements of loose coupling and ambiguity in organizations, the role of symbols in decision making and information processing, the place of myths, stories, and rituals in management, and the signifi-

cance of beliefs in the transformation of organizations (March and Olsen, 1976; Sproull, 1981), some of the distances between the properties of organizations and the properties of other social systems seem to grow smaller.

Proposing education as a possible alternative model for the design of information systems is undoubtedly disquieting. It seems possible that we know less about designing an education than we know about almost anything. At least, our confidence about it is gone. Discussions of curriculum cycle endlessly through questions of relevance without apparent resolution, and educational philosophers seem hardly less confused than we. Indeed, some recent proposals for educational reform seem to be dedicated to thinking of education as a decision support system and to tying educational activities and information rather tightly to their relevance for individual actions. It has become common to justify elements of a curriculum in terms of the improvement they provide in some specifiable activities that students will face in the future, some choices they will make.

That is not the philosophy of education -- or of life -- that we have in mind. We recall another long tradition in education and literature, one that views both education and poetry as linked loosely to a variety of ill-perceived possible future worlds and to understanding the confusions of life (Eliot, 1961; Freire, 1973). Such a vision

sees education and literature as elegant forms of idle talk, as ways in which we gain appreciation of our existence and develop our sensitivity. To describe organizational management in such terms is, of course to glorify it. It suggests that office memoranda might be viewed as forms of poetry and staff meetings as forms of theater, and we may perhaps wonder whether it would be better to admit a distinction between a divisional sales chart and a Picasso painting -- if only to assure that each may achieve its unique qualities. The dangers are real; but to a glorified view of idle talk and memoranda, we will add a romantic view of the possibilities for artistry in organizational engineering. Perhaps, with a little imagination here and there, educational philosophy and literary criticism could be used to point management information systems in the direction of a useful quality of irrelevance.

REFERENCES

Allport, G. W., and Postman, L.

1970 The Psychology of Rumor. New York: Henry Holt and Co.

Deshpande, Rohit

1981 "Action and enlightenment functions of research: comparing private- and public-sector perspectives", Knowledge: Creation, Diffusion, Utilization, 2, 317-330.

Eliot, T. S.

1961 On Poetry and Poets. New York: Noonday.

Feldman, Martha S., and March, James G.

1981 "Information in organzations as signal and symbol", Administrative Science Quarterly, 26, 171-186.

Frankenberg, R.

1957 Village on the Border. London: Cohen and West.

Freire, Paulo

1973 Education for Critical Consciousness. New York: Seabury.

Gluckman, Max

1963 "Gossip and scandal", Current Anthropology, 4, 307-316.

1968 "Psychological, sociological and anthropological explanations of witchcraft and gossip: a clarification", Man (n.s.), 3, 20-34.

Hannertz, Ulf

1967 "Gossip, networks and culture in a black American ghetto", Ethnos, 32, 35-60.

Haviland, John B.

1977 Gossip, Reputation, and Knowledge in Zinacantan. Chicago: University of Chicago Press.

Jonsson, Sten A., and Lundin, Rolf A.

1977 "Myths and wishful thinking as management tools". In Paul C. Nystrom and William h. Starbuck, eds., Prescriptive Models of Organizations (pp. 151-170). Amsterdam: North-Holland.

Lumley, Frederick E.

1925 Means of Social Control. New York: Century.

March, James G.

1973 "Model bias in social action", Review of Educational Research, 42, 413-429.

1978 "American public school administration: a short analysis", School Review, 82, 217-250.

1980 "How we talk and how we act: administrative theory and administrative life", published as the David D. Henry Lecture on Administration, University of Illinois, Champaign-Urbana, IL.

1981 "Footnotes to organizational change," Administrative Science Quarterly, 26, 563-577.

March, James G., and Olsen, Johan P.

1976 Ambiguity and Choice in Organizations. Bergen: Universitetsforlaget.

Marschak, Jacob, and Radner, Roy

1972 Economic Theory of Teams. New Haven: Yale.

Mitchell, J. C.

1956 The Yao Village. Manchester: Manchester University Press.

Paine, Robert

1967 "What is gossip about: an alternative hpyothesis", Man (n.s.), 2, 278-285.

1970 "Information communication and information-management", The Canadian Review of Sociology and Anthropology, 7, 172-188.

Rich, Robert F.

1977 "Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding", in Carol H. Weiss, ed., Using Social Research in Public Policy Making. Lexington, MA: Heath.

Rosnow, Ralph L., and Fine, Gary A.

1976 Rumor and Gossip: The Social Psychology of Hearsay. New York: Elsevier.

Shibutani, T.

1966 Improvised News: A Sociological Study of Rumor. Indianapolis: Bobbs-Merrill.

Simon, Herbert A.

1967 "Information can be managed", Think, December, 7-11.

Sproull, Lee S.

1981 "Beliefs in organizations", in Nystrom, Paul C., and Starbuck, William H., Handbook of Organizational Design (pp. 203-224). New York: Oxford.

Swanson, E. Burton

1978 "The two faces of organizational information", Accounting, Organizations and Society, 3, 237-246.

Weiss, Carol H.

1977 "Research for policy's sake: the enlightenment function of social research", Policy Analysis, 3, 531-546.

Still draft with some transcription glitches.

AN (ON-GOING) CASE STUDY IN TECHNOLOGICAL INNOVATION IN A FORMAL ORGANIZATION

Allen Newell
Computer Science Department
Carnegie-Mellon University

To appear in L. S. Sproull (ed.) <title>

29 March 1982

The ZOG research at CMU is supported by the Office of Naval Research under contract N00014-76-0874, and also partially by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 3597, monitored by the Air Force Avionics Laboratory under contract F33615-78-C-1551.

AN (ON-GOING) CASE STUDY IN TECHNOLOGICAL INNOVATION IN A FORMAL ORGANIZATION

This paper describes a technological innovation in a formal organization and speculates about its organizational consequences. Thus, it has the form of a case study. The case is solid enough, though caught in midflight, so to speak. My colleagues and I are engaged in bringing into being a sophisticated computer-based system within an operational part of the US Navy. The system is not yet fully installed, but our contact with the ultimate operational environment is already substantial. As technologists, concerned mostly with how to develop the system, we have paid little attention to the possible organizational issues surrounding this innovation. This symposium offers an opportunity to do that.

It is premature to present such a paper at a scientific gathering. My excuse is that this paper was commissioned at the last minute. It seemed both to the organizers of the symposium and to myself that there was sufficient intrinsic interest in the material, viewed as a case study, to justify presenting it. However, it is necessary to take the paper in the proper spirit and not to expect finished analysis or hard data.

The plan of the paper is straightforward. The first section presents the case in a objective way. Then follows a section with some necessary organizational and historical context. The final section asks a series of questions of interest from an organizational perspective and speculates on their answers.

1. Objective description of the case

There are three distinct components of an objective description of the case. First is the technological innovation; this is the ZOG system for processing information. Second is the real-world organization in which the application will occur; this is the USS CARL VINSON. Third is the application itself; this is the specific way that ZOG will be used aboard the USS CARL VINSON. We take up each of these in turn.

1.1. ZOG: The information system

Functionally, ZOG is a cross between a user interface, a data base and a general computational system. More specifically, ZOG is a rapid-response large-network menu-selection computer interface [RobGMN81, 1981]. The user communicates to the computer by selecting from a menu of displayed and explained choices. The computer communicates back to the user by displaying some information along with an additional menu of selections (it may also execute programs). Menu selection is a common mode of user-computer communication. What distinguishes ZOG is its full exploitation of this mode -- the unlimited use of displays and selections. This implies the existence of a large network of display *frames* (as they are called) that give information and permit selection of yet other frames. These *ZOGnets* can be tens of thousands of frames. The strength of menu selection is its self-explanatory character; its weakness is that, as expertise and familiarity

increase, selecting through a tree of options appears clumsy and time consuming. Thus, the final key aspect of ZOG is that the selection is very fast, independent of which frame in the entire network is being accessed.

Figure 1-1 shows a typical ZOG frame. Working down from the top, there is a *title*, followed by a body of *text* which displays the information offered by the frame. Next is the menu of *options*, each distinguished by a single character (1 to 3 in the figure). Each option can link to another frame, and typing the assigned character replaces the current frame with this next frame. The *option text* associated with each selection explains what the user will find. There are other menu selections on the frame, called *pads*. The row at the bottom (the *global pads*), is available on every frame and provides a basic set of actions and links for searching the ZOGnet, entering the editor, going to a root frame, etc. The others (the *local pads*) provide similar functions, but specialized to a particular subnetwork. As the pads indicate, selections can execute programs as well as link to new frames.

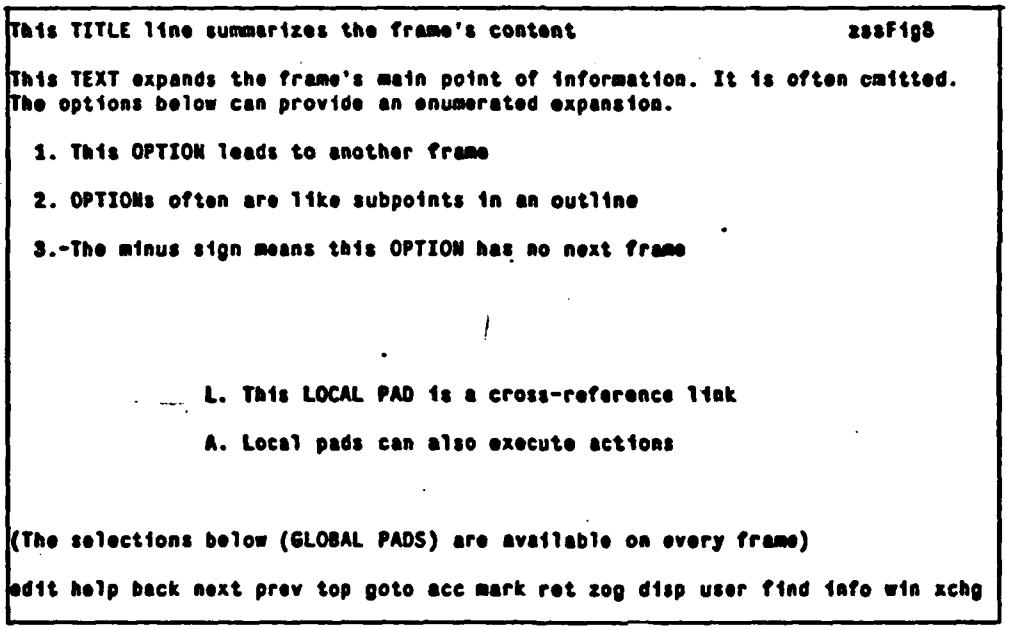


Figure 1-1: A Self-Describing, Typical ZOG Frame

To be concrete, suppose a user was faced with the frame in Figure 1-1, which was part of a ZOGnet describing the ZOG system. Wanting to know more, the user might select option 2 (by typing the character 2 on the keyboard). Immediately, the frame would be replaced by the frame shown in Figure 1-2. The user might want to go on and select another frame (e.g., at option 1 of Figure 1-2) to get more information. Alternatively, he might want to go back to the frame of Figure 1-1, which he could do by selecting the global pad, back (typing character b), or to return to the top of the net, by selecting top (typing t), etc.

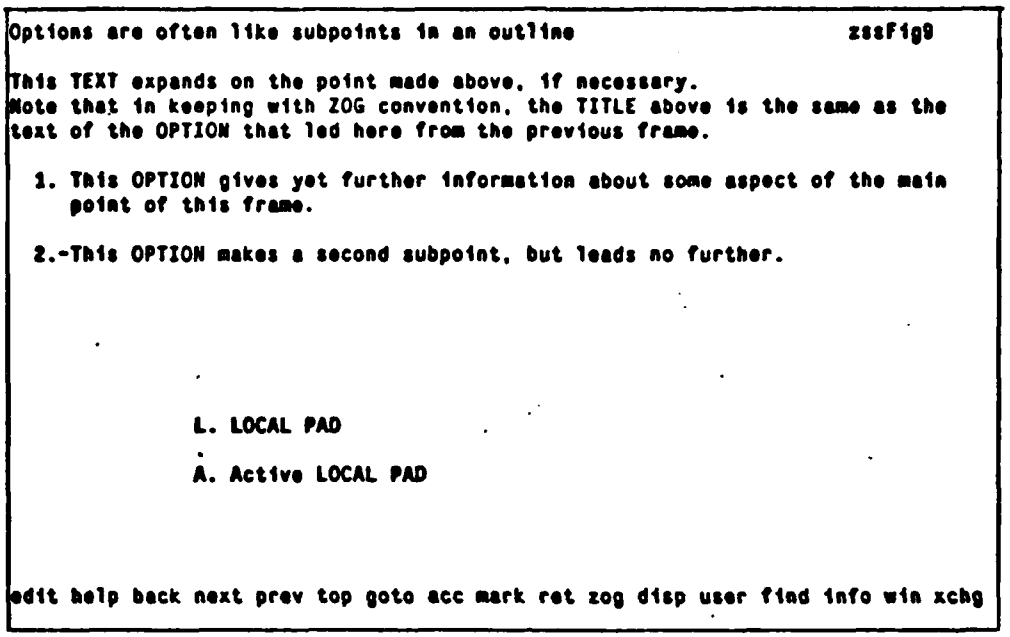


Figure 1-2: Another Example ZOG Frame

With every frame leading directly to a dozen other frames, an entire network is created. Figure 1-3 tries to give some idea of this, though it can't show all the links from global pads back to the root frame (**zog**) or other **top** frames. It does show a user-generated path though the ZOGnet (the dashed line). The **back** pad permits retreat back up this path step by step. The **return** global pad presents all of the frames of this pad as a menu of options, so the user can jump back to any prior point along the path. ZOGnets for little tasks run 50-100 frames; total ZOGnets to do a large job run into the thousands or tens of thousands. Although Figure 1-3 implies a certain locality in the ZOGnet, the time from one frame to another arbitrarily located frame is designed to be constant and rapid (i.e., less than half a second).

ZOG has other features beside the basic search and explore capability just illustrated. One is an editor which permits general editing of a frame, the creation of new frames and the linking of one frame to another. This editor (called ZED) is available at every frame via the global pad **edit**. ZOG is thus meant to be a dynamic memory structure, not just a retrieval system. A second feature is that ZOG is a communications multiplexer, which sits between the user and a general computer. It can modify messages between the user and other programs running on the computer. An important result is that a ZOG application system consists not only of a ZOGnet, but also of a collection of *agents*. These are programs evoked by the user through the ZOGnet, which can manipulate the ZOGnet and do other jobs. For instance, agents exist that create a document for a subnet in an appropriate and pleasing external form. They do this by working through *Scribe*, a modern document production system. (The present paper was created in ZOG and exists both as hard copy

Figure 1-3: Schematic ZOGnet

and as a ZOGnet.) Besides the editor and the agent capability, there is a general search capability for character strings and a statistical capability for studying the behavior of the user.

ZOG is a general purpose interface, though especially suited to situations with extensive communication to the user of data that cannot easily be algorithmically generated. In particular, data base systems that require browsing and systems that provide guidance fit well within its ambit.

ZOG has an intellectual history, though that is not of paramount importance here. It was developed by George Robertson, Don McCracken and myself [RobGMN81, 1981]. Its roots go back to an earlier system of the same name we developed in 1972. But the main impetus was a medical information processing system called PROMIS, developed by Larry Weed and his colleagues [SchuD79, 1979], which had almost all of the properties we have laid out here.

1.2. The USS CARL VINSON: The application organization

The USS CARL VINSON (CVN70) is the Navy's newest nuclear carrier. Built at Newport News, it has just been commissioned (March, 1982). It is a major ship of the line and its organization is totally dedicated to its operational mission.

The US Navy, as a military organization, is the prototype of a formal organization, with rigid lines of authority bolstered by a long tradition that supports the organizational structure. There has been some softening of the authoritarian structure in the post Viet Nam era, but it remains fundamentally intact. The US Navy is also a large bureaucracy. In this regard it is necessary to distinguish the shore-based Navy from the

Figure 1-4: Schematic ZOGnet

sea-going Navy. The latter, which includes the USS CARL VINSON, remains authoritarian, but is probably significantly less bureaucratic. This arises in part from its mission being highly operational under severe real-time constraints and in part from the extremely strong command structure.

A nuclear carrier is a large organization. Its total complement is 5600 people, 2900 of them being the crew for the ship and 2700 being the air wing. It is under the command of a captain, which is the highest rank that commands a ship (admirals command fleets and large shore organizations). Command of a nuclear carrier is the prime fleet command and its importance can be seen by the fact that the next two positions, the Executive Officer and the Operations Officer, are both also held by captains. The ship is organized hierarchically in 17 Divisions, averaging 330 persons per division, broken down into 70 Departments, averaging 80 persons per department, hence with about 4 Departments per Division, though the variance is large.

1.3. The Application

Information systems, unlike most technological devices, do not wear their use on their sleeve. Given ZOG and given the USS CARL VINSON still leaves open how ZOG will be used. As a large organization, the USS CARL VINSON has the full range of management and organizational tasks. Two specific tasks were isolated for initial attention: The maintenance and use of the SORM, the Ship's Organization and Regulations Manual; and high level planning, normally done informally, but encased here in the P&E, the Planning and Evaluation system.

The SORM is the repository of much of the standard procedures for a ship's operation. It describes what to do to accomplish all the things that have to be done. It is used in conjunction with directives that come down

from higher commands. Much of the information in the SORM also exists in the heads of experienced crew members, but the SORM still plays several critical roles. It is the court of last resort, the source of continuity, and the means of training new crew members. SORMs are large documents (the one for the USS EISENHOWER is 357 pages) which are continually changing. The SORM will exist in ZOG as a large ZOGnet (estimated 20,000 frames, with much more detail than the EISENHOWER SORM), thus completely in electronic form for updating, browsing, etc. This electronic version is to be the official version of the SORM. In addition, hardcopy versions of the SORM or parts of it will be produced from this central version.

Large organizations require some formal mechanisms to support planning, coordination and evaluation. The second management function to be provided by ZOG is support for on-line planning. Plans will exist in an integrated ZOGnet, which will be updated and modified continually as each plan is extended and changed. Exploration of plans from different perspectives will be possible, e.g., by task, by persons or by resources. Also possible will be some automatic monitoring of plans for consistency and critical events, and some propagation of completion status from frames that indicate a specific action taken to frames that summarize percent completion of higher level plans. The ZOG P&E will be aimed at the top levels of shipboard operation, at the department level and above.

Maintaining the SORM and constructing plans are independent management functions to be performed on the USS CARL VINSON, hence to be supported by ZOG. However, the SORM can be viewed as consisting of generic plans to carry out those tasks on the ship that are recurrent or sufficiently important to rate prior analysis. Thus, when dynamic planning encounters a task covered by a generic plan already in the SORM, it should incorporate that into the plan. The plan will need adaptation to local circumstances, but the generic plan is still the starting point. The consequences of this in ZOG are two-fold. First, there must be the capability to find the relevant parts of the SORM and instantiate them into a plan, thus making plan-building easier and more reliable. Second, the structure adopted for the SORM itself must be in terms of generic plans for accomplishing well-defined tasks.

To insure success, the application tasks were initially limited to two. From the standpoint of the ZOG technology, there is nothing special about the initial two. Thus, other tasks can be expected to develop and such is the long range design. One has already happened. ZOG is to be used to hold the maintenance and training manuals for the large elevators that move planes and weapons between decks. As an application, this is quite similar technically to the SORM, although the content and organization of the ZOGnet is somewhat different and there is no need to instantiate subnets (as for plans). Another application is under development: air flight management to bring the aircraft back to the carrier. This task involves both presentation of information about the current flight, and retrieval and presentation of advice about how to deal with a complex decision situation. It requires integrating modern artificial intelligence expert systems [McDermott,

1982] with ZOG.

Figure 1-5: Schematic ZOGnet

1.3.1. The shipboard system

To carry out these functions requires a system of many levels, as shown on the right side of Figure 1-6. At the bottom is the hardware system. This consists of a distributed network of thirty large personal computers (Three Rivers's Perqs on an Ethernet). On top of that comes an operating system and programming environment (SALT, a variant of the CMU SPICE development), and a programming language (Pascal). Next come the ZOG system and Scribe (a document production system), which can be likened to application-oriented higher-level languages. Then comes the specific system for performing the particular management functions (the ZOGnet format design plus agents); this is analogous to an application program. Above that comes the actual large ZOGnet, analogous to the content of a data base (and amounting ultimately to over a hundred thousand frames). Finally, at the top, there is a level dealing with the entire range of system maintenance functions: installing the system, training the ship's personnel in its use, making the system work, modifying it as experience accumulates, and gradually enhancing its functions.

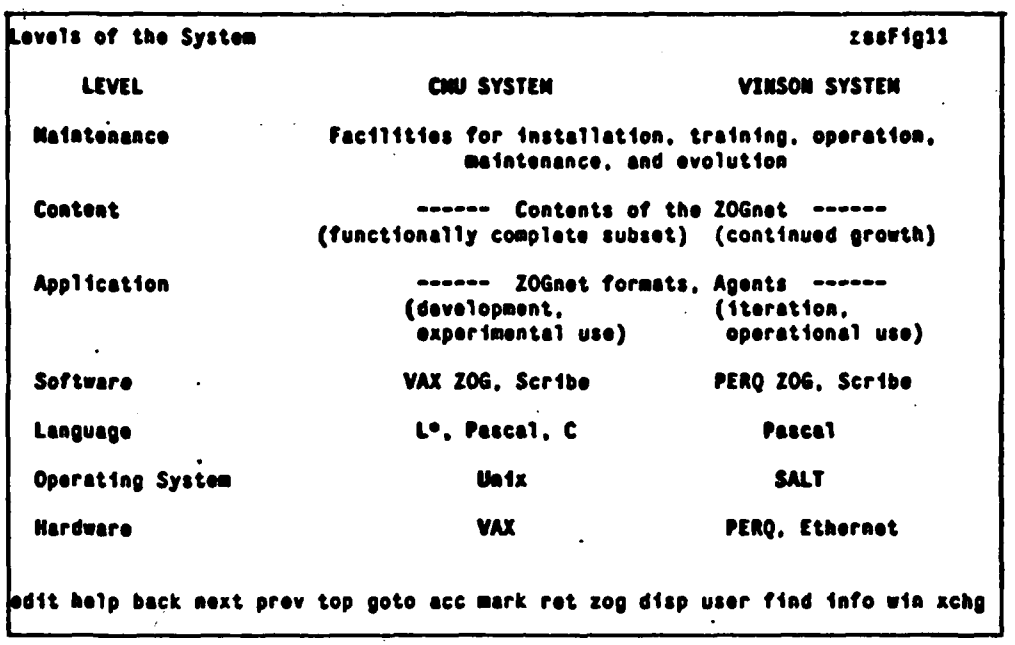


Figure 1-6: Levels of the System

The technical details of the hardware and software are not important here, except to give a flavor for the type of computational facility to be used. The computer is the Three Rivers Perq [Rosen80, 1980], which has significant computer power (1 million instructions per second), high quality bit-mapped display and substantial local disk memory (25 megabytes). Perqs operate on a high speed local area network (the 10 megabits per second Ethernet). The output will be by a Xerox multifont printer. The SPICE operating system and programming environment [Hibb81, 1981] is an experimental system being developed in the CMU Computer Science Department, to provide an advanced personal computer environment for its own use. This provides the appropriate base on which to create the distributed ZOG system. Hard copy output from the ZOGnet is provided by Scribe [Reid80, 1980], an advanced document production system developed and in use around CMU. The entire system is at the edge of the art with respect to modern interactive computation.

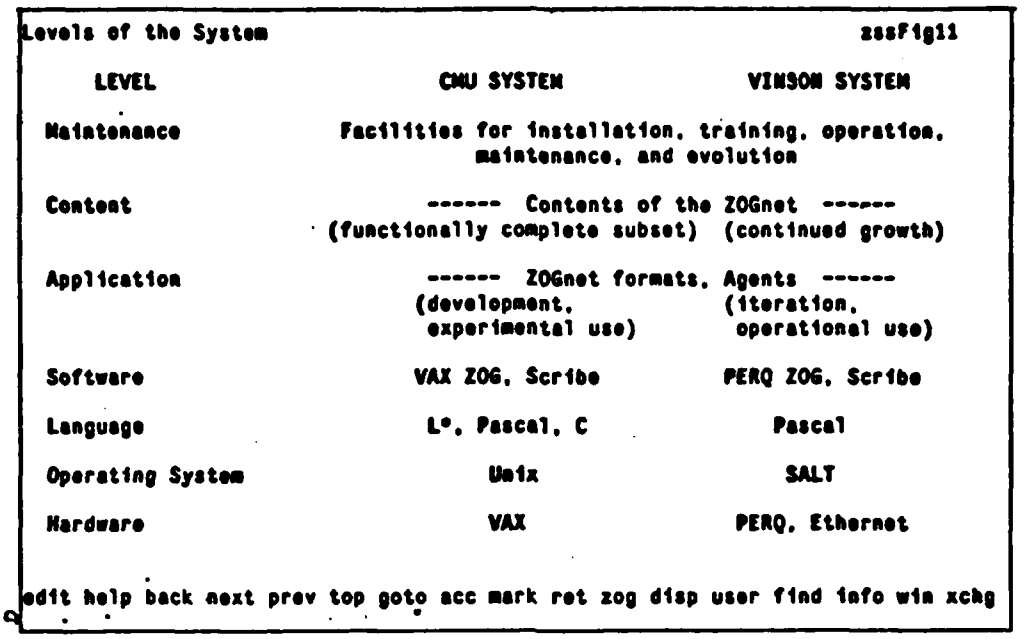


Figure 1-7: Ref to Perqs

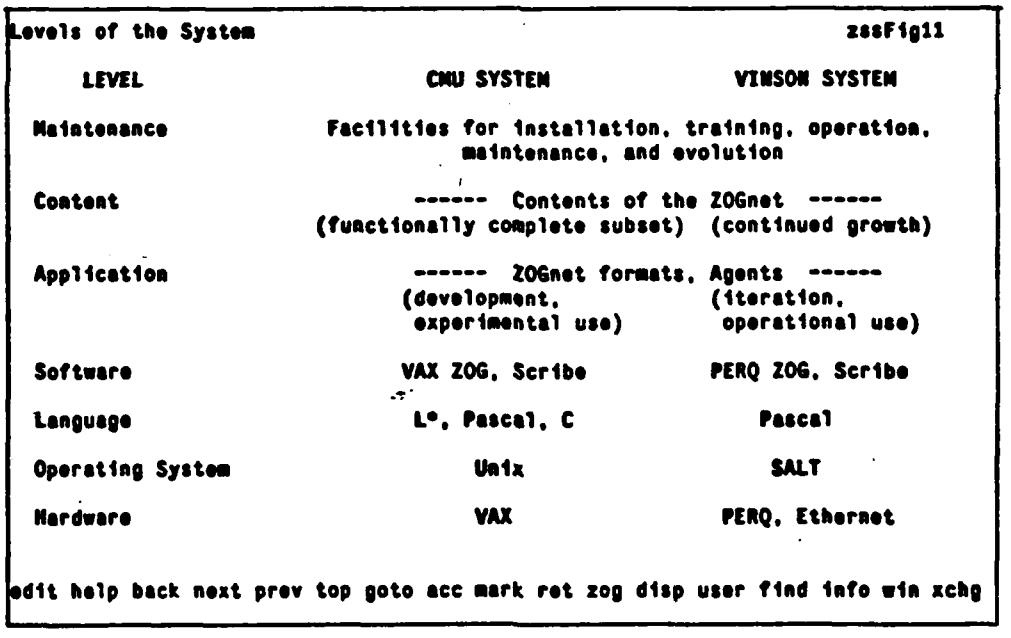


Figure 1-8: Ref to SPICE

The USS CARL VINSON must operate as a self contained unit. The system will be maintained by the ship, both for hardware and software. They will also have a modest programming capability for enhancing the

functions of the system. In this respect, the ZOGnet contains a subnet that provides a guided programming environment for creating new agents in Pascal.

2. The Organizational Context

As social scientists know, the objective description is not the whole story. In a capitalistic society there are standard models for the acquistion of new services. An organization has a need, finds a supplier and obtains a system (sometimes the supplying organization initiates by advertising). In the military, one organization develops the need, another does the procurement and a third, the operational unit, finally uses the system. In the present case the story is less standard, indicated already by the involvement of a university. This context must be given in terms of history and people. Although institutional and professional roles play an important part, there are many unique features that have substantial impact on organizational questions.

2.1. History

The story starts with government agency mission-oriented research, which operates somewhat differently than NSF and NIH. It tends to develop small groups of researchers in continuing interaction with program managers, with relatively long-term commitments and a mixture of pure and applied research. As a result, a substantial amount of loyalty and legitimacy develops. The Office of Naval Research (ONR) has done this very well ever since the late forties. In the area of computer science, Marvin Denicoff has been overseeing the research for twenty years. ZOG has been supported by ONR, since the inception of the project six years ago.

A key person in the history of the project is Captain Richard Martin, the commander of the USS CARL VINSON. A graduate of Annapolis, with an MS in Operations Research from the Naval Postgraduate College plus work towards a PhD in Applied Mathematics from UC San Diego, the Captain is a highly intelligent, technically oriented officer. He is also, as befits the position he occupies, deeply embued with the concern for command and the leadership of men. This combination of talents is not surprising, given the highly rigorous selection for the command of nuclear carriers.

The Captain was given the initial command of the USS CARL VINSON, which includes the eighteen month precommissioning period. During this period, while the ship is outfitted and its crew assembled, substantial opportunity exists to shape its basic character as an organization. The Captain has the goal of using modern computer technology to help the ship do its job. This goal extends to making the carrier a test bed -- a lead ship for introducing computer technology into the fleet. Such a goal, although not common in a line officer, is consonant with the career goals of a Navy officer. It does carry some risk in terms of innovating within a relatively tradition-bound formal organization.

Pursuing this goal, the Captain, working through Marvin Denicoff, visited many ONR research contracters examining the state of advanced technology -- in short shopping for wares. This might seem a somewhat novel course of action. However, its novelty goes back a decade, when Captain Martin set up the first F14 Air Wing. In that similar situation, as a much younger officer just out of his graduate work at UCSD, he had contacted Marv Denicoff and had also attempted to incorporate some advanced research, in the areas of training and organizational sociology.

Captain Martin first visited us in February 1980, soon after taking up his command. For us, it was just another visit, but for him it was far from casual. He was temporarily stationed at the Westinghouse Bettis Plant in Pittsburgh, where the nuclear power plants for the carriers are built, and this permitted an intensive interaction about ZOG. Part of the Captain's goal was a thorough overhaul of the way the USS CARL VINSON would handle its administrative operations. He became convinced that a ZOG-like system, implemented on a modern locally-networked collection of powerful workstations, was a feasible and exciting approach. The role of the networked powerful personal computers is essential here. Such systems are the basis on which modern office automation systems will be built in the next decade. They have substantial computing power, permitting them to be used not only for typical administrative functions, but also for advanced artificial intelligence applications. Thus, putting ZOG on such a network, not only provided ZOG, but opened the way to the full range of office and decision support functions and moved a long ways toward the general goal of the test bed.

Our reaction was, to say the least, mixed. In fact, we were looking for an appropriate driving application, having failed to find one earlier, due mostly to technological limitations. However, we were hardly in the market for an application with the large-scale, real-time, real-consequence characteristics of the sort the Captain was proposing. We were attracted by the opportunity to bring ZOG up on the next generation of computer systems (networked large personal computers). Moreover, the Captain had strong arguments for the payoff to the Navy from ONR-supported research, and good answers for most of our objections. The long-standing relations between ONR and university researchers played a crucial role here. It would have been impossible for a genuine outsider to come in and establish a basis of legitimacy for negotiations in the short time period available.

We agreed to proceed, after laying down some important constraints to help assure the chances of success, namely, that the ZOG system be brought up on top of SPICE, which provided local synergy, and that some members of the crew be stationed at CMU. The project actually got under way in July 1980, only a few months after the Captain had first made contact. Time was of the essence, of course, since a strictly limited window existed during which such a development could occur. The USS CARL VINSON was launched in March 1981 and it was scheduled to be commission in March 1982. Once fully operational the opportunity to

install a major innovation would likely be lost.

2.2. The Current Situation

The system is being jointly developed at CMU and the USS CARL VINSON. As Figure 1-6 shows, the CMU system already exists, running parallel in its structure to the proposed system to go on the ship. At the lower levels it runs on different machines (the VAX) and is built on different basic software. However, the CMU and VINSON systems share everything at the Application Level and above (plus substantial parts of the Software Level), so that the current system can be used as a development base. There are communication lines from the ship to the CMU ZOG VAX, so that personnel at the ship can work on the developing system, just as can people local to CMU.

It is difficult to pin down exactly how much total manpower is involved, because general support is being provided by both organizations. In terms of direct effort, the CMU ZOG group of about seven people, two officers of the Captain's staff, stationed at CMU, and several officers at the ship, are actively involved. Much of the ZOGnet is being built from the ship by the Navy personnel and by shipyard personnel employed by the ship. As an aside, the total manpower involved is small compared to the size of the task.

Figure 2-1 shows the current time schedule, which is tied to the schedule for commissioning and shakedown cruises of the USS CARL VINSON. A single Perq has been on the ship for some time. We have targeted for a Perq network running on the ship by Oct 82. The period from Oct 82 to Mar 83, about the time when the ship leaves on its first deployment, will go to working out the kinks in the system and modifying it to fit the needs of the ship, as borne out by actual experience. That will also be the period of transferring maintenance responsibilities away from the CMU ZOG group to another group whose resources (and temperament) are better suited to the task.

3. Questions and Speculations

The description given of the project just barely scratches the surface. The project encompasses problems that range from computer science to human factors to management science with, intertwined, issues of the research return to the scientists involved [Newell, McCracken, Robertson & Akscyn, 1982, McCracken, Robertson&Akscyn]. Discussion of all these has been suppressed and only enough has been provided to permit moving to organizational questions, which are the ones of interest to this symposium. Four questions seem to admit at least useful comments at this stage in the development of the case: (1) What are the organizational factors that bear on success? (2) What are the long term prospects for success? (3) What effect will the innovation have on the formal organization? (4) What effect will the innovation have on communication patterns?

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    :
82  |
    |- Mar ---  Ship's commissioning
    |- Apr ---  Ship on short shake down cruise (until Jun)
    :
    |- Oct ---  PERQ network running on board
83  |           Begin transfer of maintenance responsibility
    :
    |- Mar ---  Ship's deployment (approximate)
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 2-1: Project Timetable

3.1. What organizational factors bear on success?

We first focus on specific organizational factors that bear on the success of the project in the short term, i.e., in getting the system successfully installed, working and accepted. The factors to be discussed are, except for the last two, positive. It is easy enough to enumerate the standard risks of development projects -- short deadlines, understaffing, bad planning, over commitment, etc. The present project is hardly immune from these problems, but there is nothing noteworthy about its sufferings in this respect. Thus the interesting question, at least for a participant in the project, is what other factors might exist in the situation that ameliorate these standard risks. That some turn out to exist has played a part in our being willing to pursue this adventure.

Driven by the leader. The primary offsetting factor is that the project is being driven by the top man in the organization, Captain Martin. It is a commonplace of the lore of innovation that having access and approval by the top man is essential. This refers to command -- to getting decisions made promptly and favorably. However, Captain Martin's involvement goes far beyond this. He is one of the principle intellectual architects of the system at the level of the design of the ZOGnet, the key application-dependent part. The ZOGnet and its use expresses directly his management philosophy. He continuously attends to the project, even having a ZOG terminal at home, from which he both accesses and adds to the ZOGnet. Considering the immense range of his other responsibilities in organizing the ship, his attention to this project must be seen to be believed.

Cooperative definition. ZOG provides a framework within which to create a system specific to an application, by means of the conventions that rule the ZOGnet and the agents that work that net over. Above

this application-level system there is the actual content of the ZOGnet -- the formulation of the human procedures that constitutes the SORM, primitive plans in the P&E system, etc. Both these higher levels have been defined jointly by the ship and the ZOG group. Such cooperative definition is aimed directly at one of the chief stumbling blocks of technological innovation -- being designed by outsiders to the using organization. This cooperative aspect is realized in part by the existence of two officers at CMU, who have become regular members of the ZOG group, and in part by having the ship on line to the CMU development machine. The cooperative aspect is enhanced by the design being evolutionary. Thus, the cooperation has not been limited terms to a few joint conferences at design time, but continues over the course of the entire project in ways both large and small. The stumbling block of isolated design to innovational success is well known, as is the proposal to ameliorate it by cooperation. What is important is the degree to which cooperative design has been implemented in the present project.

Client origination. The project was initiated by the client, i.e., by Captain Martin and the Navy, and sold to the research group. Though a common situation in the market place, it is less frequent in the research setting, where scientists, believing their system is useful or interesting in some way, must find and convince sponsors to support it. The important operational consequence from this reversal is less tendency for the project to be oversold, hence to become committed by its creator to overly ambitious goals. One concrete example is the limitation of the initial effort to two management functions, rather than casting the system to handle the full range of management tasks.

Long-term cooperative relationship. As far as I can see, the project simply could not have occurred if there had not existed the long-term cooperative relationship between ONR and the university. Its first major role shows up in providing the Captain with legitimacy to initiate interaction with the ZOG group and present a patently outrageous proposition -- at least outrageous from a research standpoint, for it implied degrees of application far beyond what university researchers are normally prepared to do. Its second major role was permitting the university group to begin work within a few months after the initial contact. The existing contractual relations provided a base within which to operate. More important, they provided a way to uncouple the contractual process from the actual work. This involves a major degree of trust of course, because there are substantial risks involved.

The Wang system as back-up. The Captain has also obtained a Wang word-processing system for the ship, consisting of some forty terminals connected by a network with a substantial central processor. The Wang system is used for general office work. However, it has had at least four important effects. First, some initial experience has been obtained with an automatic planning system, programed into the Wang by the ship. Thus the ZOG planning system is alcady a second operational iteration. Second, however, this very activity raises a potential problem in transferring off the Wang system to the ZOG system, because the crew has gotten used

to it and has invested effort in building it. Third, the Wang system provides a back-up if the ZOG system fails to materialize. The Wang system has much less functionality, but it would still be a significant step beyond what is normally available on carriers. This possibility of a back-up system has eased the risks considerably in proceeding with the ZOG system. Fourth, the Wang system has been sufficiently efficient that it has produced the organizational slack to make available the manpower necessary to develop the ZOG system. The fixed manning tables for Navy ships implies that if you do the required job with less men, then these men become available for other tasks.

Multiple views. We now turn to a negative organizational factor. ONR is an organization whose mission is to support research. Thus, they view -- and must view -- the ZOG/VINSON project as a research effort and hence, obviously, as an experiment. On the other hand, the ship does not do research; its missions are entirely operational. For them, the system is surely experimental, in the sense uncertainty about how well it will work. However, it still lies entirely within operations (although it will be exclusively for administrative operations rather than ones involving combat subsystems). This difference of view is inherent in going directly from basic research to the field, by-passing any development stage. It continually causes difficulty in communication and negotiation. Interestingly, the problem of multiple views is classically taken to be a problem between researchers and users. Here the views of the research team are unimportant (we are comfortable with the sorts of research results and payoffs to be obtained from a field effort); the difficulty lies between the two Navy organizations.

Bureaucracy. Hanging over the whole probject is the effect of the larger bureaucracy of the shore-based Navy, which provides immense road-blocks to any innovation. The military is pro-innovation as a matter of official policy and has created a substantial organizational structure to foster the creation of new things. One effect of this is to create proper channels for innovation. This is relevant, because the present effort violates these channels, being a direct connection between a using organization (the carrier) and the researchers. Interestingly, the one other major applied effort in which I have engaged with the military, namely systems training for the Air Defense Command in the fifties, developed at Rand [Chapman, Kennedy, Newell & Biel, 1959, Kennedy, Newell, McCracken, Robertson & Akscyn, 1982&Biel], had the same direct connection between an operational unit (a brigadier general in the northwest region) and the researchers. Additional bureaucratic effects come simply from being in government. Acquisition of computers is governed by the Brooks Bill, which makes it extraordinarily difficult to obtain permission to acquire a computer system in less than two years. Indeed the funding arrangements for the project have dragged on interminably, and even at this late date threaten to keep the project from continuing.

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    :
82  |
    |- Mar ---  Ship's commissioning
    |- Apr ---  Ship on short shake down cruise (until Jun)
    :
    |- Oct ---  PERQ network running on board
83  |           Begin transfer of maintenance responsibility
    :
    |- Mar ---  Ship's deployment (approximate)
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 3-1: Project Timetable

3.2. What are the long term prospects for success?

By the long term is meant well after the ZOG/VINSON project has left the scene. Consider the range of potential outcomes. First, the system could simply not be used, after a brief initial period. This outcome has a long and honorable tradition in the military, arising in part from the separation of the using organization from the procurement organization. This outcome can stand generally for failure of the project -- we need not enumerate any of the spectacular and decisive ways the innovation could fail outright. Second, the system could work all right, but fit into a well defined niche in the existing organization. Hence it could improve the efficiency or effectiveness a bit (or even a lot), but that's all. As such it would be a success, but not an innovation. Third, the project could change the way the ship does business in some significant way. Fourth, the system could succeed, but by doing something quite different from what was planned. Either of these constitute the outcomes of interest to organizational scientists. Fifth, perhaps, there is the possibility that the innovation could spread to other elements of the fleet, although that is beyond the horizon of the present paper.

I list below some of the important factors that seem to bear on which of these outcomes will occur. Despite the array of organizational factors listed above, most of the factors enumerated here are not organizational but depend on other aspects of the situation, namely underlying technology, features of the application and general culture. In any event, this list gives my estimate in order of importance.

Technology: Reliability and speed. The most important factor, I believe, is the basic performance parameters of the system. The most important of these are reliability and speed. The functionality of systems is almost always useful, by design. That the present system is being designed cooperatively with the using

organization increases that probability substantially. However, if the system is sufficiently unreliable, it will simply not be used. Even if the system is thrust crosswise into the operation of the ship, so its use seems required, ways will be found around it. Likewise, if the system is too slow by any appreciable factor, it will likely be subverted and retired. In the present case, the software has a long history and is quite robust. However, the hardware (the Perqs) is new and its performance, especially under sea-going conditions, is untested. In addition the low level software (SPICE) is also new.

Psychology: Ease of learning. The next most important factor is how easy the system will be to learn. This seems more important than how easy the system actually is to use. Once users get familiar and relatively expert with a system, we know they tolerate all manner of intricate and non-functional ceremonials -- society would be hard put to operate, if this were not the case. But if the barrier to becoming familiar is too great, the opportunity is provided for all the other forces against system success to come to bear. Historically, computer systems have been hard to learn, not so much from any intrinsic complexity, but from the degree to which the interface with the user is arbitrary, unforgiving and full of meaningless incantations. The evidence about ZOG in this respect is fairly encouraging, including the direct evidence from its use by Navy and shipyard personnel during the last year over remote connections. These people seemed to have little difficulty learning ZOG. However, this problem cannot be taken as solved.

Application: Rigidity of functionality. The next factor is the rigidity of the system. The real needs of the organization and the real environment are different than envisioned at design time. Even the technique of cooperative design, which is specifically aimed at counteracting this, only helps a little, because the operational personnel are hardly more able than the scientists to foresee how an innovative system will really operate. Thus, unless the system has considerable built-in flexibility, it will not in fact perform well enough to survive. This factor, like the ones above, leads mostly to the system ultimately not being used in any significant role in the organization. Computer systems have a mixed history on flexibility. On the one hand their programmability has been touted as making them the most flexible of technical devices. Though true, it has proven remarkably difficult to realize that flexibility in actual user systems. ZOG has a designed strategy for dealing with this, which we call *semi-automatic* operation. It is the same strategy that makes text editing systems so useful -- though sometimes awkwardly, they can always be employed by the user to get his text transformed appropriately. Again, we shall have to wait to see whether the flexibility of ZOG manifests itself sufficiently.

Culture: Infiltration of computer in society. The potential is always present for resistance of an organization to a technological innovation on general procedural grounds, i.e., that it conflicts with the organization's existing ways, especially if it introduces new types of demands. Computers certainly have been in this situation. However, although there is only anecdotal evidence, the situation in the Navy may be in the

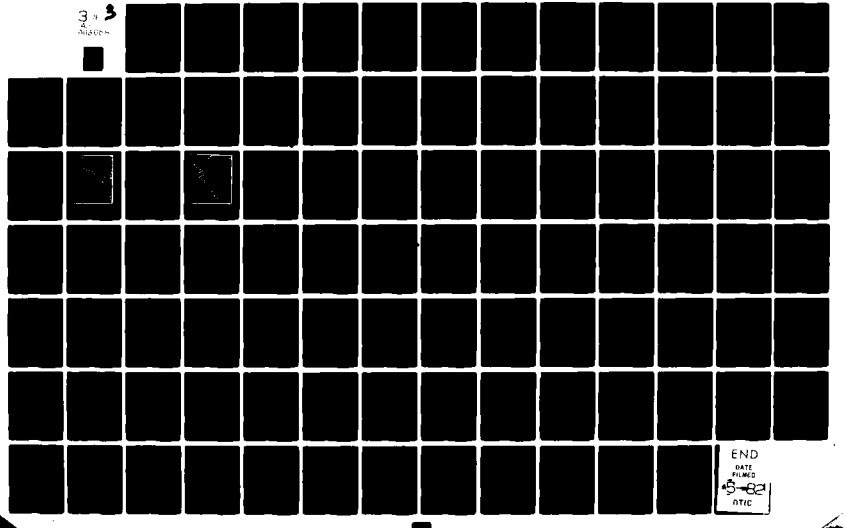
opposite situation. Most immediately, the fleet (as opposed to the shore-based Navy) is relatively starved for computers [Toedt81, 1981] -- despite the contrary inference from the shipboard world being highly technical. The reasons for this are complex, but are based on a combination of bureaucracy and on the needs for shipboard systems to be combat hardened. One continually runs across stories in which officers get their own personal computers in order to do this or that. However apocryphal, it indicates an attitude toward the computer. Even more, it indicates that the rise of the personal computer is creating an external environment in which many people have contact with the use of computers and do not see them as an intrusion, but as a normal part of the scene.

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    :
82  |
    |- Mar ---  Ship's commissioning
    |- Apr ---  Ship on short shake down cruise (until Jun)
    :
    |- Oct ---  PERQ network running on board
83  |           Begin transfer of maintenance responsibility
    :
    |- Mar ---  Ship's deployment (approximate)
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 3-2: Project Timetable

Organization: Larger embedding bureaucracy. Bureaucracy was already mentioned in connection with the short term prospects. It is no doubt a sign of its pervasive influence that it shows up again in connection with the long term. An activity not supported in regular channels is continually open to tranformation by the larger organization. Though use of the system can possibly proceed for short periods of time, simply as a component of the ship's operation, almost any event will bring it back into interaction with the larger organization -- attempting to expand the number of workstations is a typical and likely possibility. A major (and proper) stance of the larger organization towards an innovation is to ask about its larger implications. Indeed, a mark of its success is that it have implications for the entire fleet or at least all carriers. Conversely, if it has no implications, then it clearly cannot have been worthwhile for the USS CARL VINSON, for a ship is not an individual, but a member of a class of ships (even though simultaneously it is seen as having a unique qualitative character). But as soon as the system becomes a symbol within the larger issue, its fate is no longer tied to local issues, such as technical efficiency or even overall organizational usefulness, but again becomes political.

AD-A113 658 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF SOCIAL SCIENCE F/G 9/2
SYMPOSIUM ON INFORMATION PROCESSING IN ORGANIZATIONS.(U)
APR 82 L S SPROULL, P D LARKEY N00014-82-G-0004
UNCLASSIFIED NL



END
DATE FILMED
5-82
DTIC

Organization: Personnel turnover. The second organizational factor is the short tours of duty of Navy personnel. It is really a special aspect of the larger bureaucracy. Tours typically last two years and their duration is not related to the tasks being performed. As a result, no one stays around to finish a task and ensure its ultimate success. The Navy is full of lore about the sad fate of innovations under the shift of commands. In fact, there is some flexibility in this system and it is relevant to us. The commanding officer of a nuclear carrier does have a longer tour (about four years), an exception due, I believe, to Admiral Rickover. But already key people have left the project to move on to other billets. Of course, the Navy has more general turnover problems, with men leaving the service. But this latter affects the project more through the issues of ease of learning the system and the level of sophistication of the user community.

Technology: The larger SPICE system. The final factor is again technological, namely that the system is actually placing in operation an advanced interactive system for office automation. This ties into the Captain's vision of the USS CARL VINSON as a test bed for moving research in computer science and artificial intelligence onto the ship. From the present concerns, it increases the probability that new functions will be found for the system that turn out to be more important than the original ones. This can be viewed as a planned increase in functionality; but I think it is more properly viewed from an organizational standpoint as a serendipitous shifting of the nature of the innovation. Whether this will occur is, of course, problematical and depends on some rather uncertain issues about how the system will be utilized aboard the carrier, whether stringently or with considerable freedom, and what sorts of relations develop between the carrier and the ONR-supported research community. This latter is strongly affected by the problems of turnover. Already there are additional technical imports on the ship from elsewhere in the ONR community and one new AI project has been spawned at CMU on planning flights that directly involves the USS CARL VINSON.

3.3. What effect will it have on the formal organization?

The Navy has a time honored organization. Though its texture has apparently shifted somewhat in the post Viet Nam era to accomodate a change in the ethos of its personnel, the structure itself remains intact. A major structural impact on the formal organization is not to be expected from the present innovation. Even if computation is destined to have large impacts, as have other technologies such as the telephone, it would happen through long accumulation. The present system would play an infinitesimal role in the early stages. However, it is reasonable to ask whether the ZOG/VINSON system will shift or stress the structure in any particular direction. Two possibilities are worth brief examination: authority and the rationalization of tasks.

Erosion of authority. The ZOG/VINSON system implies the presence of essentially public data bases that contain much of the procedures and plans for the organization. It seems at least plausible that such

accessibility might erode authority by eroding the base of special knowledge on which authority in part depends. There is no way in knowing yet, of course, but it seems unlikely. In fact, I have been impressed with the way the crew mixes authoritarian and cooperative regimes and seems to keep them straight. Less anecdotally, there is considerable evidence in the work of Rob Kling [Kling, 1980] on the introduction of computers into formal (but not military) organizations that computers work to confirm existing structure. Computers can always be used in many modes and those in control initially witting or unwittingly shape the modes of use to reinforce the status quo. For instance, counterbalancing egalitarian aspects deriving from widespread knowledge of plans is the increased ability of the authority structure to monitor plan completion.

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    :
82  |
    |- Mar ---  Ship's commissioning
    |- Apr ---  Ship on short shake down cruise (until Jun)
    :
    |- Oct ---  PERQ network running on board
83  |           Begin transfer of maintenance responsibility
    :
    |- Mar ---  Ship's deployment (approximate)
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 3-3: Project Timetable

Rationalization of tasks. The SORM produced in ZOG is a big rationalized procedure net. It appears to be much more rationalized than any existing SORM. Developing the SORM within a hierarchical top-down structure with consistent conventions works strongly in this direction. Even stronger is the desire to use the SORM as an operational basis for constructing plans by copying them over and instantiating them to local situation. Behind this is the Captain's belief in the importance of such rational task structure. This can be taken as a form of Taylorism, and such an identification used to predict difficulties with its implementation -- that it will rapidly be subverted because it will be too restricting. Much depends on the style of implementation -- how tightly behavior is expected to conform and how key is the role of planning, with adaptation of the plan to the current situation. Much also depends on whether the SORM becomes viewed as common property or imposed structure. I do not think we have any strong handle on predicting this yet, as most of the assimilation of the system to the operational style of the ship is still to occur. It is important to note, however, that in the context of the onboard nuclear power plant, the Navy has successfully implemented very strict procedures.

3.4. What effect will it have on communication?

The best candidate for radical change is the communication pattern. By this is meant the actual pattern of messages between individuals throughout the day, taking the system to be a communication channel. (The system might itself come to be taken as an actor in the communication net, though we will not explore the possibility here.) The system must be taken to include all the functions to be performed by the network of distributed computers, not just those concerning the SORM and P&E, the initial developments.

Electronic mail. The basis for the prediction is the effects produced by existing electronic mail systems. The main novel feature of electronic mail is the decoupling of sending and receiving (like ordinary postal service, but unlike the telephone) with instantaneous transmission (like the telephone, but unlike ordinary mail). Other features are also important, e.g., long term indexed storage of messages, broadcasting conversation-like interchange if both parties are linked simultaneously, and (ultimately) integration into other automatic processing. A shift away from personal communication to the use of the mail system can be expected. However, given the formality of vertical communication on board ship, important side consequences for organizational structure would seem unlikely. More important would be the possibility that role of electronic mail would expand quite generally beyond task information. Unfortunately, little is known, I suspect, about existing patterns of communication against which to measure any effects. This shift, of course, is not unique to the ship, but will be occurring wherever extensive electronic mail systems are installed. If the ship is especially interesting, it is because it is an operational, rather than an office, environment.

What could effect the outcome. Several features of the ship could limit any impact of such a new communication media. First is the finite size of the ship and the guarantee that people can be found within it via, say, the intercom. The size of a carrier might nullify this, but it is still a real factor, I believe. Second, the number of work stations is only thirty. This may be enough for the planned functions, but be nowhere near enough for general communication. Thus, not until there are 300 work stations available will there be an important effect on the overall communication pattern. Finally, the efficacy of existing technology should not be forgotten, namely, the little green book carried by every officer to record personally important facts and plans. If this remains a preferred mode of operation, electronic mail might be only another channel to get information from senders to be stored in the green book of receivers, and there would be very little effect except some increase in efficiency.

4. Conclusion

I have tried to describe a technological innovation in a formal organization, one of particular interest to this audience because the innovation involves information flow. The innovation must be taken as still potential, though it is already embedded in the USS CARL VINSON and is being applied in minor ways. For example,

the ZOG SORM has now become the official document for how the ship will operate; and (hardcopy) plans for the acceptance test prior to commissioning were produced via the system. In addition to describing the case objectively, I have given a modest amount of the historical and organizational context. Finally, I have touched on a number of organizational effects and considerations surrounding the case.

It is tempting to claim this innovation is unique. I have certainly given the story that flavor. But when I compare this case with other analogous situations I have known, it seems to be true of all real situations that they seem unique. This, of-course, is just the institutionalist's refrain and the analyst's lament. However, in favor of the analyst (whose side I am really on), it must be said that all the unique factors I have listed seemed important precisely because we have a general model of their effects. We believe access to the top leader in an authoritation organization is crucial for success, because of our general (and overly simple) view of control in such organizations plus the general (and overly simple) belief that organizations are not hospitable to innovations from outside. Thus, the particular mix is unique, but the ingredients come from familiar stock.

That is the good news. The bad news for organizational analysis, it seems to me, is that the factors that seem to affect the case -- its long-term success or its evolution -- are catholic in their nature. Some are technological, some are cultural, some are psychological and some are organizational. They all interact to determine what will happen. Better expressed, they all pose intersecting constraints within which the destiny of this case will work itself out.

Finally, I must reiterate the disclaimer at the beginning. I did not promise you a real analysis. Indeed, in my present role of system developer, I could not have provided it, even if the time had been available and the spirit willing. Thus I present the case for what it is -- an on-going case study that you may find interesting and useful for speculation.

5. Acknowledgements

This paper draws heavily on [Newell, McCracken, Robertson & Akscyn, 1982, McCracken, Robertson&Akscyn], including using its figures. I want to acknowledge its authors, Don McCracken, Rob Akscyn and George Robertson, who are also the CMU principals on the ZOG/VINSON Project, along with myself. I especially want to acknowledge Capt. Richard Martin, who has been a prime initiator and mover on the project. As is always the case in extended research projects, many people, both from the Navy and from CMU, have contributed to ZOG and to the ZOG/VINSON project in many different ways and I also want to acknowledge their efforts.

```
80  |- Feb ---  Begin discussions with Capt. Martin
    :
    |- Jul ---  Official start of CMU involvement
    :
81  |
    |- Aug ---  Ship's crew moves on board, with 4 remote CMU terminals
    |- Sep ---  Initial on-board use of P&E, remote to CMU
    |
    |- Nov ---  First PERQs to ship: stand-alone, running subset of ZOG
    :
82  |
    |- Mar ---  Ship's commissioning
    |- Apr ---  Ship on short shake down cruise (until Jun)
    :
    |- Oct ---  PERQ network running on board
83  |           Begin transfer of maintenance responsibility
    :
    |- Mar ---  Ship's deployment (approximate)
    |           Enhancements to on-board system complete
    |           End of direct CMU involvement
```

Figure 5-1: Project Timetable

References

Chapman, R. L., Kennedy, J. L., Newell, A. & Biel, W. C. The Systems Research Laboratory's air defense experiments. *Managment Science*, 1959, *5*, 250-269.

Hibbard, P. G. Research in Personal Computing at Carnegie-Mellon University. *Computer Science Research Review 1980-1981, Carnegie-Mellon University*, 1981.

Kling, R. Social analysis of computing: Theoretical perspectives in recent empirical research. *Computer Surveys*, 1980, *12*, 61-110.

McDermott, J. R1: A rule-based configurer of computer systems. *Artificial Intelligence*, 1982. (in press).

Newell, A., McCracken, D., Robertson, G. & Akscyn, R. ZOG and the USS CARL VINSON. *Computer Science Research Review 1980-81*, 1982.

Reid, B. *Scribe: A Document Specification Language and its Compiler.* Doctoral dissertation, Computer Science Department, Carnegie-Mellon University, October 1980.

Robertson, G., McCracken, D., and Newell, A. The ZOG approach to man-machine communication. *International Journal of Man-Machine Studies*, 1981, *14*, 461-488.

Rosen, B. PERQ: A commercially available personal scientific computer. In *COMPCON 80 Spring*. IEEE-CS, 1980.

Schultz, J.R., and Davis, L. The technology of Promis. *Proceedings of the IEEE*, September 1979, *67*(9), 1237-1244.

Toedt, Lt. D.C. Shipboard information handling tools. *United States Naval Institute Proceedings*, January 1981, *107*(1), 104-106.

Table of Contents

List of Figures

INFORMATION AND AMBIGUITY IN ORGANIZATIONAL CHANGE

J. Rounds

Mental Retardation Research Center
University of California, Los Angeles

Prepared for the Carnegie-Mellon University Symposium
on Information Processing in Organizations

October 1981

INFORMATION AND AMBIGUITY IN ORGANIZATIONAL CHANGE

Whether as organizational scientists or as citizens, we all know that large, bureaucratic organizations are tenaciously resistant to change, even when confronted with information that they are failing to meet their goals. And yet, it is an obvious historical truism that sometimes organizations do change, and sometimes they change in radical, fundamental ways. The question, then, is why the defense mechanisms that normally protect organizations against fundamental change sometimes fail. More specifically, for the present paper the question is this: under what circumstances (if any), and through what processes, does the acquisition of information contribute to fundamental organizational change?

By fundamental change I mean change so complete that the belief structures underlying the goals and technologies of the organization have been altered in significant ways,so that the actors' understanding of what they are trying to do, how they are trying to do it, and why their methods should work are all changed. Fundamental changes in organizations are comparable in scope to the shifts in scientific paradigms described by Thomas Kuhn (1970; 1977). In the terms of political scientists they are revolutions rather than reforms (e.g., Garner 1977:6).

I do not mean that an organization must metamorphose into an entirely different type of organization in order to have undergone a fundamental change. Prisons, for instance, have been able to change radically three times in American history and still remain recognizable as prisons. They retained the general goal of the control of crime, and incarceration as a basic technique, but revolutionized their understanding of the nature of crime and criminals, their fundamental strategies and practices for the use of incarceration to combat crime, and the rules under which they accepted and released prisoners (Rounds 1979).

Fundamental change, then, does not consist of the incremental changes or refinements in technology that are a constant feature of virtually all organizations, however stable. Instead I am dealing with relatively infrequent changes that affect the deep structure of organizational beliefs (Sproull 1981) and the practices constructed upon them. Such beliefs are central to the fundamental "definition of the situation" (March and Simon 1958:154-155) or "closed system of variables" (Simon 1976:82,83) that gives the organization coherence and sense.

Simon and March have described how these closed system are formed, and how they function in organizations (Simon 1976, Ch. 5; March and Simon 1958, Ch. 6). Organizations, they argued, provide a social mechanism for the achievement of rationality far beyond that possible for any individual, given the cognitive limitations of the human mind. One way they do this is by organizing and simplifying the informational environment in which members of the organization make decisions. The rich complexity of that environment is divided into two classes. One--the "closed system"--consists of all those types and sources of information defined as germane to the situation in which the organization must decide and act. The other class is the "empty world" outside the closed system, a realm in which nothing occurs that is of relevance to organizational decisions. Thus, to members of the organization acting in their roles as its decisionmakers, there is no information except that to which attention is prescribed by the closed system.

The closed system is conveyed to new members of the organization not in the form of abstract precepts, but largely as routinized sets of practice, "performance programs" which specify proper procedures under given circumstances (March and Simon 1958:141-148). Since the closed system is a summation of the experience of many persons, it is in principle impossible for any individual to understand

fully why all of the programs of the organization are the way they are, or the basis for each of the multitude of past decisions that fixed them in their present form. At any time the validity of the bulk of the closed system must simply be taken for granted.

The closed system, or definition of the situation, then constitutes "a simplified, screened and biassed model of the objective situation" (March and Simon 1958:154-155). It organizes the recognition and use of information by members of the organization by "selecting (1) particular values as criteria for later decisions, (2) particular items of empirical knowledge as relevant to later decisions, (3) particular behavior alternatives as the only ones needing consideration for later choice" (Simon 1976:97-98).

In the process of formation of its closed system an organization comes to resemble an onion. At the center are the base assumptions, the fundamental beliefs about the nature of relevant phenomena upon which all else is constructed. Around this core are wrapped layers consisting of the specification of the central values of the organization, and of the boundary between the closed system and the empty exterior world. Outside these layers are others made of the performance programs that constitute the action elements of the organization's technology. The most basic programs are layered closest to the core, while those that are least central to the organization's identity are located at the periphery. Once the onion is formed, the stability of its elements increases the closer they are located to the core. Growth--or incremental change--takes place in the outermost layers, and seldom penetrates far inward.[1]

A major reason for this is the buffering provided to the inner layers by those closest to the surface. The programs in the outer layers were fixed by earlier decisions (Simon 1976:97-98) based upon criteria and information that are then lost through the process of "uncertainty absorption" (March and Simon

1958:165). That is, only the decision itself, and the resultant programs, are passed on to the next generation of actors, but not the original premises for the decision itself. In terms of the metaphor, the premises remain in their original layer of the onion, while the decision provides the basis for a new layer. This aids in expansion of the rationality of the actors by eliminating basic information that must be gathered and decisions that must be made before action can be taken; but at the same time that it makes it unnecessary to recheck fundamentals before each decision, it also makes it extremely difficult to do so should the need or desire arise. Once the system has attained any degree of complexity, the members of the organization can no longer easily recover the chain of decisions, and their premises, which led to the current performance programs. To anthropomorphize, the organization has forgotten how and why it came to be the way it is. "Genesis amnesia" has set in (Bourdieu 1977:79); fundamentals of the system have receded into the unconscious foundations of practice.

This does not mean, however, that those fundamentals have in fact disappeared, or lost their power to organize present behavior. Like the unseen hand of history that makes puppets of us all, they are there in the programs that guide the organization's routine. History has become misrecognized as nature (Bourdieu 1977:78-79). That is, arbitrary decisions made on the basis of limited information have become perceived as revelations of the nature of reality, through the process of uncertainty absorption. The decisions have declared how the world is; their assertions are sacred, in the sense of being unquestionable (Moore and Myerhoff 1977:3).

Thus we see how the internal layers of the closed system onion are buffered from direct scrutiny by organizing the information environment of the actors, to ensure that their attention will be focused upon the program layers at the periphery. Information and decisions are required only where there is

uncertainty, and the process of forming a closed system is to move the fundamental belief structure of the organization into the realm of bedrock certainty. At the same time the large surface of the outermost layer provides a more-than-ample supply of uncertainty to occupy the puzzle-solving skills of the actors, uncertainty which also serves to account for failure without resorting to questioning of the fundamental beliefs.

I began by asking how the acquisition of information contributes to those occasional episodes of fundamental change in organizations. A traditional model of pure rationality in organizational behavior would hold that the connection between information and change should be very strong and direct. Since organizations are presumed to be goal oriented in such models, when they receive information demonstrating that their activities are failing to meet those goals they should change rapidly to more promising activities.

Sometimes organizations do seem to behave that way in regard to decisions about allocating effort among the possible activities specified by their closed systems. That is, where the definition of the situation includes programs for trial and evaluation of alternative activities, then information does result in change through a process that would meet the popular criteria for rationality. These changes occur in the peripheral layers of the onion, and their occurrence does not disturb the inner serenity of the closed system.

The role of information in fundamental change, however, is more problematic. Fundamental change by definition involves change in the core elements of the onion, and the closed system makes it unlikely that members of the organization will recognize information that brings those basic beliefs into question. There are at least three forces that inhibit such recognition: (1) attention is diverted to the periphery and away from the core by the performance programs; (2)in many cases the actors are unaware of the underlying premises of their

programs, and so are incapable of recognizing the information as relevant to their behavior at all; (3) where basic beliefs are conscious, they are perceived as matter-of-fact observations of the nature of reality, so that denials of their validity are viewed as bizarre or even psychotic.

It is possible, of course, that an accumulation of information at the periphery could add up to a critique of the fundamentals, but this also seems unlikely. The information attended to under the direction of the performance programs makes sense only in the context given by the closed system, which specifies what the decision issue is, what information is relevant, and how that information should be interpreted. There is a deep chasm between use of information in this manner and use of information to question the very premises under which the data were recognized, collected and interpreted.

Kuhn has described the corresponding problem in scientific paradigms, which are the closed systems of those loose organizations he calls scientific communities. Anomalous data--those which contradict the predictions of the paradigm--are always present, but in times of normalcy scientists do not perceive those data as challenges to the validity of the paradigm. Rather, they treat them only as particularly difficult puzzles, and remain confident that eventually they will be solved within the existing system. Under a stable paradigm the object of routine scientific research "is to solve a puzzle for whose very existence the validity of the paradigm must be assumed. Failure to achieve a solution discredits only the scientist and not the theory" (1970:80).

Similarly, while in organizations specific solutions adopted within the bounds of a closed system may be evaluated as failures, there is no way to reach a similar judgement about the definition of the problems themselves or the principles for seeking solutions without stepping outside the system. However, the closed system denies that relevant information exists outside, and directs

attention away from sources of information that might contradict its fundamental assumptions. If the closed system of an organization is really such a powerful mechanism for organizing the informational environment and attention patterns of its members, then how can its hold ever be broken? How is fundamental change ever possible?

Two Models of Information and Change

This paper will present a case study which should clarify the problems raised in abstract above. The case involves change in the organizational system which delivers services to the mentally retarded in California. Since early in this century these services were concentrated in a single type of organization, the large, residential state hospital operating on a medical/psychiatric model of treatment. In the 1960s the system began to shift toward provision of services through "community-based" facilities. Over the past twenty years the patient populations of the state hospitals have shrunk considerably, while there has been an even greater rise in the numbers of retarded receiving community-based services. This is widely regarded as a revolutionary change in the service system.

There is a popular story, told by both laymen and social scientists, which holds that this change was due to the acquisition of information in the late 1950s and early 1960s which demonstrated that the hospital technology was actually the cause of many of the symptoms of retardation that it was supposed to treat. Bizarre behavior, dependency and the absence of progress in development of life skills were discovered to be the results of living in "total institutions" rather than of innate disabilities. A movement formed around the community services alternative, which was based on a radically different view of retardation and the role of the retarded in community life. The movement has had

considerable success in developing community-based services, but has been less successful in its goal of "deinstitutionalization"--that is, of shutting down the old state hospitals and ending their counterproductive effects upon their inmates. This is because of the usual inertia of large bureaucracies and because the power structure of the system remains under the control of traditionalists whose personal careers and power are dependent upon the continuation of the hospital system. Thus, the process of change over the past twenty years has been a confrontation between those acting rationally on the basis of reliable, valid information, and those acting from personal motives, in defiance of the negative information concerning the hospital system, to preserve the traditional structures.

Similar tales of "speaking truth to power" are told of many other kinds of organizations as well. Indeed, this may be the most pervasive story form in the literature of applied social science. Translated into the metaphor of the organizational onion, this theory of change holds that the death of the onion begins at the core and moves outward to the periphery. That is, change begins with rejection (by some individual or small group) of the most fundamental assumptions of the closed system. Even after this is accomplished, though, the performance programs located in the peripheral layers continue to run. In fact, they prove far harder to conquer than did the core. Thus, the change agents must work their way outward, destroying programs which no longer have any excuse for their existence.

This model is intuitively pleasing, and yet it is difficult to see how such a thing could happen, given the arguments of Simon and March about the nature of closed systems. Certainly it is possible for an individual to disagree with any fundamental assumption of an organization--especially an outsider who has not been socialized into the closed system. But this is possible only if it is recognized what those fundamental assumptions are, and I have argued that this

usually is not the case. The core is buffered effectively by the outer layers, so that the fundamental assumptions are deeply embedded in the unconscious foundations of practice. Those basic assumptions which are conscious and can be cited by members of the organization are typically general beliefs about humans, society or nature that are widely shared throughout the society, so it is unlikely that deviant opinions on such matters will find a constituency with sufficient power to force change from outside.

In sum, the theory that change begins with an individual or small group rejecting the fundamentals of the organization's definition of the situation appears to violate the psychological foundations of Simon and March's explanation of the genesis and function of closed systems. Such systems overcome the limited rationality of the individual mind; yet this theory of change is dependent for initiation of the change process upon a heroic feat of individual rationality in which some person perceives the essential nature and fundamental assumptions of the existing closed system, correctly recognizes and interprets information falsifying those premises, and creates an alternative system. When an organization and its closed system are operating effectively, this seems an inherently unlikely event. The theory that the onion dies from the center outward seems to require different psychological principles for the explanation of change than were used for the explanation of how the system was originally constructed.

It would be more consistent with the principle of bounded rationality to hypothesize that our organizational onion dies in the opposite direction, from the periphery toward the core. This would make it unnecessary to hypothesize any exceptional act of individual rationality to set the process in motion. Since the closed system of an organization has been built by a process stretching over a long period, and has provided such elaborate buffering for its own core,

this approach suggests that an elaborate process of unlearning the system must precede any direct attack on that core, and the reconceptualization of the system from new fundamentals.

A preliminary abstract statement of that process may clarify the argument made below in the context of the case study. The process begins at the periphery, with information indicating failure to meet some acknowledged, limited goal of the system. This is interpreted as indicating some specific flaw in design or implementation has retarded the success of an otherwise viable system. The actors may perceive this flaw as central to the technology, but from the viewpoint of the outside observer it will be seen to reside in peripheral elements of the system. The action taken is not radical, but is rather a conservative attempt to preserve the system with the least possible increment of change.

The corrective measures taken by the actors to correct this flaw have the unanticipated consequence of raising some new ambiguity about the system, this time located a layer closer to the core of the onion. Another flaw is perceived, and further corrective action is taken, which again leads to a heightened sense of ambiguity. An elaborate feedback loop develops, which results in stripping away the layers of the onion one by one. As each is removed more of the buffering which protects the core is lost. At last the core stands naked, and is itself perceived as ambiguous. At this point the actors have been released from the controlling power of the closed system, and they may begin constructing a new definition of the situation without the need for heroic acts of individual rationality. Thus, in this model the ground is prepared for a radical reconceptualization of the field through a transitional process in which conservative attempts to protect the system have the unanticipated consequence of undermining it. Where the model of change that depicts the onion as dying from the core toward the periphery sees a transitional process marked by

conflict between two well-articulated and mutually exclusive systems, this model sees a transitional process marked by rising dissatisfaction, confusion, dissension, eclecticism and drift, all summed into a spiraling sense of technological ambiguity.

In the remainder of the paper I will present a more thorough version of the standard account of information and change in the California mental retardation services system that is based on the "core-to-periphery" model of change. I will then present evidence that the actual pattern of information and change has followed the "periphery-to-core" model, and will elaborate on some of the mechanisms by which information is used in breaking down the buffering of the core. Obviously in the present format it is impossible to cover a twenty-year-long process of change in much detail, and so I have concentrated on aspects that distinguish readily between the two models. More complete documentation of this case history must await completion of a monograph now in preparation.

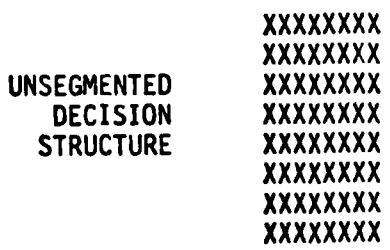
Change in California Mental Retardation Services

In 1960 almost all retarded persons receiving state services in California were congregated in five state hospitals, most of which were far removed from the major population centers that had formerly been the homes of the bulk of their residents. A substantial percentage of those residents remained in the institutions until their deaths. By 1981 the population of the state hospitals had shrunk from around 13,000 to 8,200, and more than 60,000 developmentally disabled citizens were receiving "community-based" services through twenty-one regional centers located throughout the state.

Several other concurrent changes in the system are usually pointed to a integral parts of the reform. These include the themes of democratization, advocacy, equity and individualization.

Democratization of the system is often referred to as "giving the consumers a voice" in decisionmaking, with the term "consumers" expanded to include the parents of the retarded as well as the retarded themselves. This broadening of the base of participation in decisionmaking has taken several forms, including establishment of a consumer-dominated state policy council, boards of directors for each community-based service, legal requirements for parental involvement in developing program plans for their children, and so forth. Besides the general argument of the right of concerned citizens to participate in government, this change is frequently justified by the claim that "parents know best" what services their children need, since they have the most intimate day-to-day experience with them.

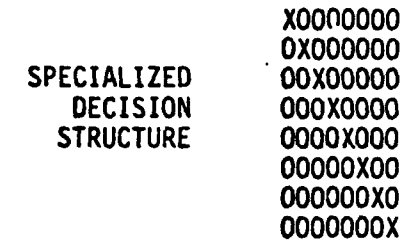
For present purposes it will be helpful to describe this change more abstractly. March and Olsen (1976, Chs. 2 and 3) have described three basic forms or ideal types of decision structures. The simplest is the unsegmented structure, in which all members of the organization are presumed to have the right and the competence to participate in all decisions. This structure can be represented by an array in which the columns signify decision opportunities (i.e., all the types of decisions made or situations in which decisions are made) and the rows represent the various members (or classes of members) of the organization, while "X"s represent access to participation in the decisionmaking process. Thus:



Stable unsegmented decision structures are typically associated with simple

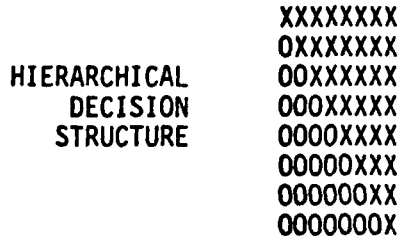
technologies and norms of egalitarianism.

There are two types of segmented decision structures. In a specialized decision structure access to specific decisions or classes of decisions is limited to specific individuals or classes of members who are presumed to have some unique competence relative to the type of decision. Thus:



Stable specialized decision structures are typically associated with a complex and varied technology. There is a high degree of professionalization, but general administration is placed in the hands of professional managers rather than with the professionals who perform the core technological activities of the organization.

A somewhat more complex type of segmented decision structure is the hierarchical system, in which individuals or classes at the upper levels of the organization are presumed to possess competence to participate in all decisions, while at the lower levels individuals or classes have access only to an increasingly limited number of decision opportunities. Thus:



Stable hierarchical structures are likely to be associated with a less complex or varied technology than in the case of specialized systems. There is high professionalization, but here the core professionals also hold the general

managerial positions (or, at least, core and managerial positions are all held by members of a single profession). The convergence of professional socialization and the organizational system of social control encourages a high degree of stability and predictability in decisionmaking.

Prior to the 1960s the California mental retardation system--the state hospital system--fit most closely this hierarchical type of decision structure. Physicians--in particular, psychiatrists--were not only the core professionals of the system, they were also its general managers. By 1981, though, the system had moved in significant ways toward an unsegmented structure, in which the base of participation in decisionmaking had considerably broadened, particularly at the top and bottom of the decision opportunity spectrum (i.e., in large-scale strategic system decisions and in the day-to-day specific decisions on which clients would get what services). Obviously the system at no time was a perfect example of either of these simplistic ideal types, but the contrast of pure forms does capture important differences between the states of the system at the beginning and present stage of the change process.

The imperfect achievement of an unsegmented decision structure is reflected in the strong emphasis in the system on advocacy, defined as the protection of the rights of the retarded. In part this refers to the general civil rights of the retarded, which are said to have been violated by lifelong segregation in a state hospital, but advocacy also focuses on ensuring that the client receives the services to which he or she is entitled. It is presumed that the service bureaucracy will inevitably withhold required services if constant vigilance is not maintained, and action taken either through the formal appeals procedure or through court suits. This is an active area, with specially-funded legal-aid-type agencies as well as area advocacy boards and clients' advocates on the staff of each regional center. By contrast, in 1960 a client refused service had

little recourse. In the opinion of many, advocacy is an outgrowth of the broader American civil rights movement, and lies at the very core of the system reform.

While advocacy protects the client against the system, the equity issue raises the problem of competition among clients for the limited service funds available. Typically this refers to an equitable distribution of resources between clients served in the remaining state hospitals and those in community-based services, but it also refers to a generalized expectation of equality in the receipt of benefits and to such special issues as the allocation of funds among the twenty-one regional centers.

The final theme commonly cited as central to the reformed system is "individualization", meaning that programming should be designed or packaged to meet the special requirements of each client. This is taken to contrast with the pigeonholing of clients into a few fixed programming categories said to have been typical of the traditional state hospital system. As this implies, programming in the system has become far more complex and diversified. The service menu is much larger, and the necessity for choice among alternative services much more frequent.

In sum, over the past twenty years California mental retardation services are said to have changed dramatically, with the central features being the substitution of community-based for hospital-based services, democratization of decisionmaking, formalized protection of the civil and service rights of clients, and provision of more diversified and individualized services.

As noted above, there is a widely accepted--and quite spurious--story about how the changes of the past twenty years came about. This popular history has been repeated to me by actors in all roles in the system, and appears in professional analyses as well.[2] According to this story, change in mental retardation services began as a direct result of information revealing the

failure--indeed, the intrinsic evil--of the old state hospitals. Halpern, et.al. (1978) cited studies appearing in 1954, 1957, 1962 and 1966 as having contributed to the recognition that "to a large extent the dependence, apathy, hopelessness, and bizarre behavior commonly observed in institutions were actually fostered by the features of institutional life." "Since institutions were seen as 'bad' influences on clients, some observers demanded that they be closed and abandoned and that clients be returned to the care of their communities." "As a result, the decade of the 1960s saw a massive decline in the number of residents residing in mental hospitals and, to a lesser extent, those residing in institutions for the retarded" (1978:1).

Taylor and Bogdan (1980:210-211) attributed the movement to information that "challenged the legitimacy of the state schools for the retarded", criticism that "has come from a range of respected sources, has received widespread publicity, and has called into question the legitimacy of the very existence of institutions as a form of service organization." These criticisms emanated from "the consumer movement", "extension of the civil rights movement", "federal legislation" and "social science perspectives", all of which followed upon a series of "exposés" beginning around 1965.

In sum, the story holds that the acquisition of negative information about the effects of hospitalization stimulated the formation of a political alliance of parents and friends of the retarded with certain concerned professionals, who have exerted pressure for change in the face of resistance by the traditional entrenched professionals. As a result of successful organizing much of the original reform program has been achieved, though deinstitutionalization is still far from completed. Thus, change began when an established closed system was challenged at its most fundamental level, and was confronted with an alternative--and radically different-system.

I have already argued that this core-to-periphery account of change is logically implausible. It is also inaccurate history. Based upon an intensive review over the past fifteen months of all recoverable primary source documents left by those central to the initiation of the change process, three key features of that process will be shown to contradict the standard account and to support a periphery-to-core model of change instead. First, information leading to rejection of the fundamental value of the hospital system was not central to the initiation of expanded community-based services in the 1960s. Second, the information and initial innovations that began the change process were in peripheral, not core, elements of the system. Third, since the initiation of the process there has been a growing sense of ambiguity concerning the system which has reached progressively closer to the core.

Unfortunately, the value of this case study is ultimately limited by the fact that the change process is not yet complete. Thus, it cannot demonstrate that a radically new reconceptualization of the system can arise once the buffering around the core of the closed system has been stripped away, since nothing of the sort has yet happened in the California mental retardation services system. Nevertheless, the case study does present a clear basis for differentiating between the two models of information and change.

Initiation of the Change Process

Given the seemingly endless chains of events which constitute social processes, it is a bootless enterprise to seek a specific date on which the change process can meaningfully be said to have begun. Nevertheless, 1963 was at least a critical year in which the perception of a need for change became widely accepted in California and actions specifically intended to foster change were taken. In that year Governor Edmund G. Brown appointed a special Study

Commission on Mental Retardation, charged with preparation of a plan for redevelopment of the service system. In the state legislature the Assembly Subcommittee on Mental Health Services was established to perform a similar function.[3] These two committees, the California Council for Retarded Children (the major statewide parents' group) and the staff of the Department of Mental Hygiene (which administered the hospital system) became the major forces in initiation of the change process.

Why were the committees formed in 1963, and not before or after? Without meaning to suggest any historical inevitability to the events, they seem to have resulted from a unique combination of opportunities moving downward from the federal government and pressures moving upward within the state system. President Kennedy had a retarded sister and so a personal interest in the topic. Early in his administration he appointed a special panel on mental retardation, which issued a report in 1962 recommending greater attention to community-based alternatives to institutional services. Federal legislation providing funding for state-level planning for mental retardation services followed in 1963, and the Study Commission was formed in anticipation of the availability of these funds for its work.

However, formation of the committees was not a simple response to another opportunity to draw federal funds into the state. The intensity of the attention turned to the issue by key decisionmakers was stimulated largely by information recently acquired that heralded an impending crisis in the existing institutional system. Figure A shows the growth in the number of patients in the state hospitals in this century. The period prior to 1963 had seen a rapid acceleration in the size of the hospital system, with a rise of about 73 percent in the previous ten years. This was reflected in the costs of the system, as shown in Table 1: capital investment in the hospitals had risen 46 percent just from 1955 to 1962,

FIGURE A

Patients in California State Hospitals for the Mentally Retarded, 1902 - 1979

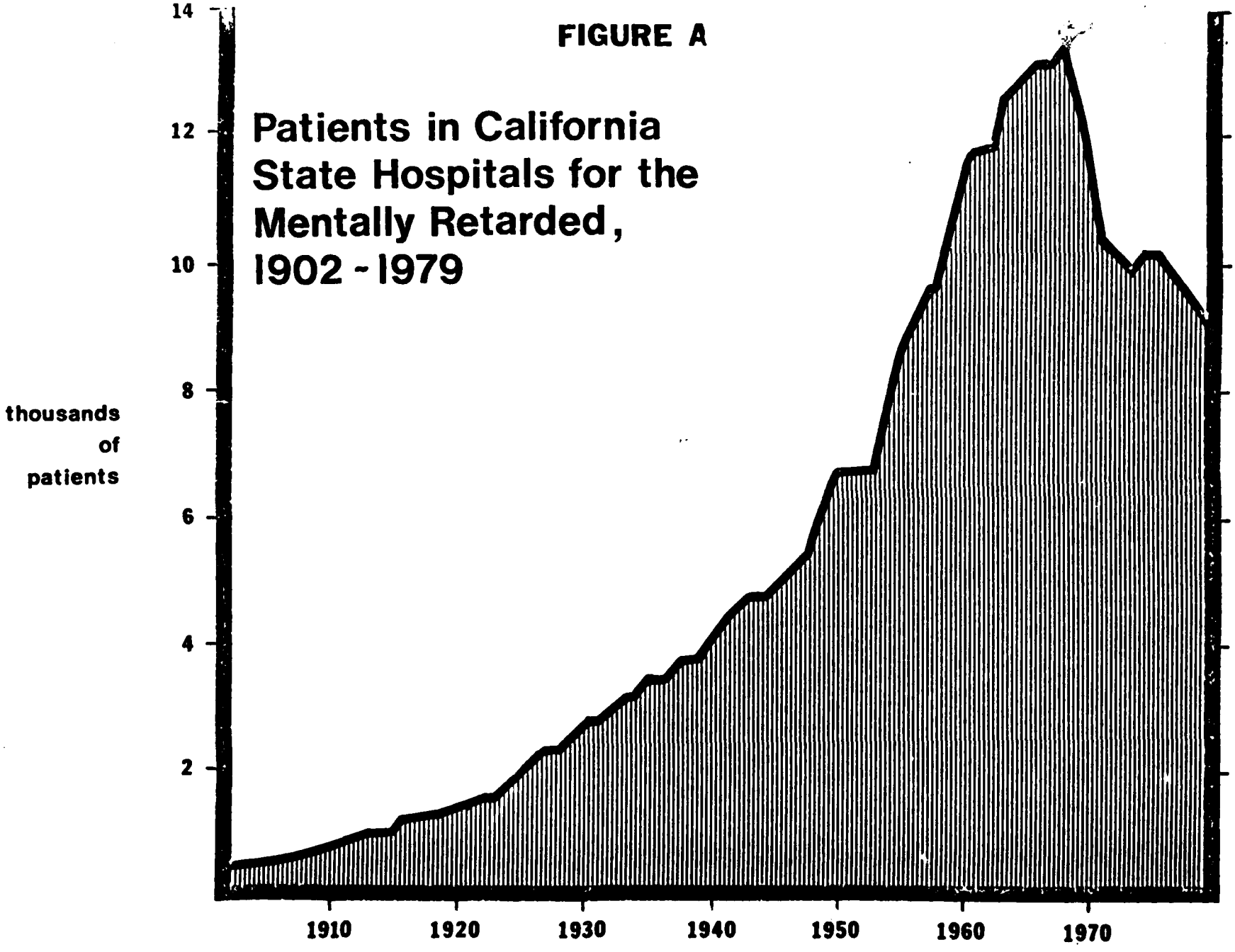
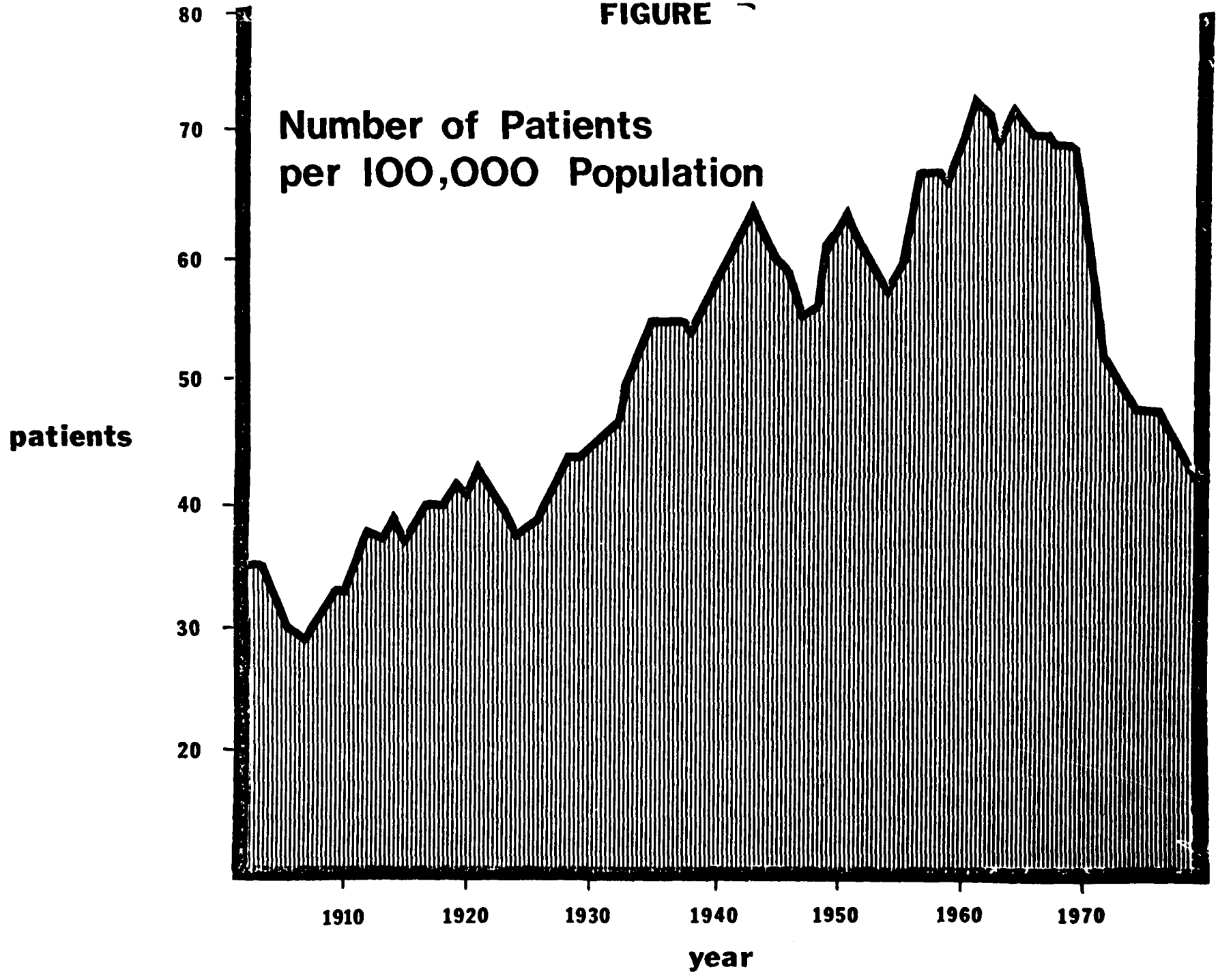


TABLE 1

CALIFORNIA STATE HOSPITAL
SUPPORT EXPENDITURES AND CAPITAL INVESTMENT,
1955-1964

year	total capital investment (in millions)	annual support expenditures (in millions)
1955	$50.3	$15.2
1956	55.4	17.9
1957	56.7	21.5
1958	60.6	24.7
1959	65.4	29.2
1960	66.9	31.9
1961	69.8	34.2
1962	70.5	39.6
1963	73.0	43.0
1964	73.6	44.7

FIGURE


Number of Patients
per 100,000 Population
patients
80
70
60
50
40
30
20
1910
1920
1930
1940
1950
1960
1970
year

while annual operating expenditures were up by over 160 percent.

This dramatic increase in the size of the system was due in large part to the rapid growth in the general population of the state in the post-war years. (The annual increase in the 1940s and 50s averaged 4.18 percent, as compared with only 1.44 percent in the 1970s.) However, this was not the sole cause. Figure B shows that the general trend of the system was toward a constant increase in the rate of hospitalization per 100,000 state population. (The figure shows numerous peaks and valleys because the system, which was virtually always operating at full capacity, grew in large increments whenever new facilities were added, while the state population grew steadily.)

Two factors were major contributors to this rise in the rate of hospitalization. First, medical advancements had increased the life expectancy of the more severely disabled members of the hospital population, which resulted in a lower turnover rate for the available hospital beds. Thus, a stable intake rate would result in a steadily increasing hospitalization rate.

The second factor was the continuous problem of the large waiting lists for admission to the state hospitals. Even at their peak the hospitals were never the residences of more than a few percent of the retarded citizens of the state. The majority of these persons never sought admission, but there were always more prospective patients than beds available. From 1955, when statistics were first given in the annual reports of the Department of Mental Hygiene, to 1963, there was an average of over 1900 persons formally enrolled on the waiting lists. Moreover, large expansions of the hospital system made only the most temporary reductions in the waiting lists: as soon as the waiting time for admission was reduced significantly, more of the 97 percent of the retarded who were not in institutions would be rushed onto the waiting lists by parents who apparently had previously despaired of receiving service.

Members of the legislature and officials of the Department of Mental Hygiene

were constantly reminded of the waiting list problem by their constituents who were parents (or other relatives and firends) of the retarded persons who were awaiting hospitalization. From first application to admission frequently took up to three years, and during that period the financial and emotional demands of caring for a severely disabled child were having a devastating effect on many families. As the parents' groups had become more effectively organized through the 1950s the pressures on administrators and legislators to solve this problem had intensified.

Until around 1963, though, the response to these pressures followed the same performance program it had for some decades: add more hospital facilities and free some existing beds for the most severely disabled persons on the waiting list by discharging the most mildly retarded and well-adjusted of the hospital residents into the community aftercare system. But around 1963 a new strategy emerged with relative suddenness and considerable drama.

Before detailing this innovation, a methodological note is in order. It is not the intent of this paper to analyze the psychology of information use in innovation--or, more specifically, how the solution to the waiting list problem was "discovered" in the mind(s) of one or many of the actors in this process. The question is certainly one of great interest, but is beyond the boundaries of my professional competence. The following account of the emergence of the solution is, then, focused at the level of analysis of the organization system, and uses only a very simplistic notion of the underlying psychology of discovery. The source materials do not suggest to me that that discovery process was characterized by any logical chain of thought deduced from basic factual information. Rather, the production of the solution seems to have been a process of "pattern recognition" (Ziman 1978), in which a number of elements present simultaneously in the environment were creatively joined in a manner which was

suspectible to public presentation as a matter of logical inevitability. In other words, the particular form of the solution was largely a matter of historical accident, in spite of making perfectly good sense. Thus, the following account is deliberately lacking a tight sequential structure that might imply an order I do not perceive in the actual events.

The emergence of the system innovations appearing in the few years around 1963 are seen here as stimulated by the coincidence of at least three major forms of information reaching key decisionmakers: the intensified information on the problems of families with children on the waiting list, which was being provided by increasingly powerful and militant parents' organizations; information on the financial implications of continuing past policies on facility expansion provided by the research staff of the Department of Mental Hygiene; and information on possible programmatic alternatives, supplied initially through the reports of the President's Panel and other sources. It will be noted that I have not included information about the deleterious effects of institutionalization on this list. I will return to that issue later, after reviewing what do appear to have been the major influences on the beginnings of the change process.

The development of the parents' groups and their increasing pressure on legislators and administrators has been noted, and will not be described in any greater detail here. Information on the consequences of continuing the standard program for expansion of the hospital system crystallized early in 1963, in a bulletin prepared by the the Biostatistics Section of the Department of Mental Hygiene (DMH 1963a). That paper analyzed trends in state hospital growth, taking the current count plus the waiting list as indicative of the total beds needed. Projecting five years ahead, the study concluded that maintaining that level of service (a hospitalization rate of 80 per 100,000) in the face of the rising state population would necessitate an increase of capacity from the

then-current 12,372 beds to 16,170 beds by 1968.

This study evoked considerable consternation, particularly among the legislators who would be asked to appropriate money for this latest expansion of the system. The Assembly Subcommittee on Mental Health Services translated these numbers into cost projections, concluding that construction costs alone would run some $47.6 million by 1968 (about 65 percent of the total capital investment in the system in 1963), and would increase the annual operating expenses of the system by approximately one-fourth (ASMHS 1964a:3). But even this huge expenditure was not perceived as the end of the road. The Department of Mental Hygiene report also included a survey of other states, which showed that even where a hospitalization rate twice that of California had been achieved there was still a problem with large waiting lists (DMH 1963a: Table 12). The dismal conclusion, according to the Subcommittee, was that even if the 4,000 additional beds were constructed "as this is being done, we will accumulate a new 'waiting list' and by 1968 we will have to build again!" (ASMHS 1964a:3).

While it is clear that the legislators thus sensed that past growth patterns could not be sustained indefinitely, this realization in itself does not account for why it was at this particular point in the growth curve that action was finally taken. As shown in the figures and tables above, the rapid acceleration in size and costs of the system was a real phenomenon, but there is no objective sense in which it reached crisis proportions in 1963 rather than at some other time. The same may be said for the pressures emanating from the parents' groups.

The additional element that served to crystallize action at that particular time appears to have been the legitimation of an alternative to the standard program by an external authority--the President's Panel--which focused attention

on a previously underused alternative, community-based services. The initial report of the President's Panel (1962) placed considerable stress on the development of community-based services for those retarded persons who did not require institutionalization, and argued that many current hospital patients could be better served by such programming. By early in 1963 efforts were underway to secure information that would confirm this possibility in California.

In August 1963 the Department of Mental Hygiene released a "Survey of Patient Needs for Residential Care and Assistance" (DMH 1963b) covering both hospital residents and those on the waiting lists. That study concluded that only 36.3 percent of the hospital patients and 39.6 percent of the waiting list cases actually required residential services in a state hospital (Table 2). These estimates were picked up by the Subcommittee on Mental Health Services, which issued its first report, "A Preliminary Proposal to Eliminate 'Waiting Lists' for State Hospitals for the Mentally Retarded", in June 1964. That proposal declared that "Logic and economy dictate that a higher priority be given to relocating the inappropriately hospitalized than to constructing new facilities" (ASMHS 1964a:4).

The Subcommittee members felt that the information they had collected--through reports and a series of public hearings--indicated that the surprisingly high percentage of innappropriate hospitalizations was due to "a basic flaw in the way State responsibility is structured" (ASMHS 1964a:4). A major incentive causing some unknown but substantial percentage of parents to place their children on the waiting lists was not their inability to cope with needs of the children at home, but the large financial demands involved in doing so. Under existing law admission to a state hospital was the sole access to public assumption of this financial burden or any part of it. Thus, many parents who were willing (or even anxious) to keep their child at home were forced to seek

TABLE 2

AUGUST 1963 ESTIMATES OF NEEDS FOR HOSPITALIZATION
AND OTHER FORMS OF SERVICE

type of care required	percent of hospital patients	percent of waiting list	percent of total
hospitalization	36.3	39.6	36.7
24-hour nursing home care	30.3	6.9	27.5
foster home care	28.6	21.6	27.8
home care	4.7	31.4	7.9
other	0.1	0.5	0.1

[source: DMH 1963b:33]

hospitalization in order to avoid bankruptcy.

A follow-up survey of the parents of persons on the waiting list revealed that about half would not place their child in a state hospital if financial assistance were available for alternative forms of service (ASMHA 1964c). Another study found that community-based services already existed, but were seriously underutilized because their costs were beyond the private means of most parents --in spite of being less expensive than the per capita costs of the state hospitals (ASMHA 1964b). The Subcommittee concluded that the perennial problem of pressure for endless expansion of the hospital system could be solved simply by correcting the flaw which gave parents a financial incentive to use the hospitals instead of other, less costly forms of service.

This plan enjoyed consensual support among the parents' organizations, the Department of Mental Hygiene, and the Study Commission, as well as the Subcommittee. The final reports of the latter two groups (SCMR 1965; ASMHS 1965) both stressed the allocation of state support for community-based services, and adopted a recommendation of the President's Panel that access to community services should be coordinated through a "one-door" referral agency able both to connect clients with generic services (such as welfare) for which they were eligible and to fund other needed services directly where generic services were unavailable. This agreement culminated in the passage in 1965 of Assembly Bill 691, establishing regional centers in San Francisco and Los Angeles which were to serve as prototypes for the present system of twenty-one such centers throughout the state. The change process had begun.

Anti-institutionalism and Conservativism in the Change Process

The central question raised by this paper is how change can emerge from a situation that discourages attention to information that might falsify the

fundamental assumptions of the closed system. In this light, the important issue to be raised about the case study is whether negative information about the established technology was known to the initiators of the change process, and whether they interpreted such information as challenging the basic technology.

In considering the first question we are necessarily limited to observing references made to critiques of institutions by the key actors. Since such references are usually made in the context of justifying a proposal made or an action taken, this is obviously different from knowing exactly what material each actor had read or otherwise been exposed to. However, such ideal knowledge seems unattainable, and relying on explicit references at least seem preferable to assuming that the material must have been a major influence upon the thinking of the actors simply because it existed in professional journals.

In all the primary source material reviewed from this period (which included every recoverable published document of the Subcommittee, the Study Commission, and the Department of Mental Hygiene), only one explicit reference to articles on the negative effects of institutionalization was found. A few other statements were discovered which suggested a familiarity with that material (e.g., Begab 1963:25), but only once were actual citations or quotations given. This occurrence was in the final report of the Subcommittee (ASMHS 1965:19-21), which stated "There is a massive body of professional knowledge and opinion that clearly documents the effects of institutionalization in large facilities remote from the natural family and the normal community". A long quotation followed, which was described as "typical of research findings on this matter". It concluded

> The data of this study supports the point of view that the institution is not as desirable a setting as is an adequate home environment for the development of trainable mentally retarded children (Cain and Levine 1961).

The Subcommittee also cited (but did not quote) four other studies, published in 1944, 1956, 1960 and 1961. Summing up the implications of this material, the

report argued that

> There is no doubt that we will better serve the interests of trainable retarded children, and perhaps even other more severely retarded children, if we redesign our system to provide home care help and other community-based alternatives to state hospital placement for those families who desire such choices (1965:20; emphasis in original).

Two points should be noted: first, information on the effects of institutionalization seems a very minor theme in these reports. Only in this single instance is it given explicit attention, and even here it received only brief notice compared with the other arguments for the development of community-based services. Furthermore, this seems to be the first time the issue was raised in the reports of the Subcommittee, which had been presenting the economic justification for over a year. Thus, while information on the negative effects of institutionalization was known and cited by the initiators of the change process, there is no basis for claiming that it was a significant source of the impetus toward change.

Second, it is clear in the citations above that the authors were interpreting that information as relating to the effects of institutionalization upon only some of the retarded: the "trainables", or less severely handicapped. It is implied that hospitalization should remaim the proper response for those unable to profit by community services. Immediately after the passages quoted above, the Subcommittee moved to dispel any doubts on this point:

> The committee wishes to make it quite clear that its purpose is to expand the choices available to families. This does not exclude the state hospital choice. For families who prefer state hospital services... and for families whose children may require state hospital services, state hospitals should be available (1965:20-21; emphasis in original).

This was not an idiosyncratic position: I have not found a single call for complete elimination of the state hospitals until the early 1970s. In the early and middle 1960s the state hospitals were seen as "one of several" among the needed resources for care of the retarded (President's Panel 1962:134). It was

argued that "We cannot eliminate the hospitals since there will always be need for a medical facility to care for the severely retarded" (Philips 1963:21). The director of a parents' group, Hope for Retarded Children of Santa Clara County, stated flatly "the State hospital has got to exist. It is the answer" (testimony of Patricia Hobbs before Senate Fact Finding Committee on Multi-Purpose Centers for the Retarded, January 13, 1964; see also testimony of R.A. O'Reilly and M.D. Stein before Study Commission, January 16, 1964).

Most telling of all, the initiation of the regional center system was actually accompanied by a continuing expansion of the hospital system. After receiving the report of the Study Commission, Governor Brown delivered a special message on mental retardation services to the legislature, offering a seven-point plan including the first two regional centers, a $300,000 allocation for planning a new state hospital in the San Jose area, and another $300,000 for planning a 500-bed "training and rehabilitation" center. This followed a bond issue of more than $5,000,000 for "high-priority construction and building modernization projects" for the state hospitals at the end of 1964. "The Governor said that the need for hospital construction remains...despite the development of alternative methods of care through private nursing care placement from the hospital and the work to be done by the regional diagnostic and counseling centers" (Mental Health Progress 6(1):12, 6(6):10-11 [1965]). The superintendent of one of the state hospitals, who is generally regarded among the most important figures in the development of alternative programs, was able to state confidently that "As a result of present nationwide planning, the establishment of many new hospitals for the mentally retarded can be expected" (Tarjan, Eyman and Dingman 1966:535). Figure A shows that hospital populations in fact continued to rise until 1968.

It is thus clear that the people who initiated state-supported, community-

based services in the middle 1960s did not interpret the negative information on the effects of institutionalization as proof that the hospital technology was fundamentally wrong. In fact,they interpreted it as actually supporting well-established principles of the technology. There had been a brief period early in the century when the eugenics scare had led specialists in the field to call for the institutionalization of all retarded persons, however mild their disability. The financial implications of such a goal ensured that no state made a serious effort to meet it, but in any case the principle itself was very short-lived among the experts. By 1920 such major proponents of institutions as Walter Fernald and Henry Goddard had been convinced by their own studies of retarded persons in the community that most could make a successful adjustment with no danger to themselves or society (Davies 1930:92).

Since the 1920s it has been an accepted tenet of mental retardation services that only a minority of retarded persons should be placed in state hospitals. Opinions on the exact percentage have varied over time and among experts, but the actual rate of hospitalization in California changed little over time. As shown in Table 3, if we assume that the percentage of the general population who are retarded has remained stable at around three percent[4], then the portion of retarded citizens of California who were residing in its state hospitals increased only from 2.0 percent in 1940 to about 2.4 percent in 1963. In practice over 97 percent of the retarded have always been regarded as capable of life in the community. Thus, the assertion that many retarded could better be served in the community hardly seemed radical to the hospital establishment. When the President's Panel declared that "Institutional care should be restricted to those whose specific needs can be met best by this type of service" (1962:135), no one disagreed.

A simple diagnosis was available which posed no challenge at all to the

TABLE 3

PERCENT OF RETARDED POPULATION
IN STATE HOSPITALS, 1940-1963

year	estimated state population	estimated retarded population	state hospital population	percent of retarded hospitalized
1940	6,950,000	208,500	4,076	2.0
1945	9,344,000	280,320	4,999	1.8
1950	10,643,000	319,290	6,703	2.1
1955	13,004,000	390,120	8,510	2.2
1960	15,863,000	475,890	11,556	2.4
1963	17,530,000	525,900	12,686	2.4

value of hospitalization for the severely retarded: a good system had gone awry because of an administrative flaw that encouraged violation of the principles of the technology, causing persons to be hospitalized who should have remained in the community. What was rejected in the 1960s was not the hospital technology, but the assumption that the system could continue expanding at a faster rate than the state population--a problem which was declared to have been the result only of "inappropriate hospitalizations". As that popular term implied, there was little doubt that appropriate hospitalizations also occurred.

Belying the image in the standard account of a confrontation between the conservative protectors of the hospitals and the progressive proponents of community services, creation of the regional center system enjoyed strong support among the hospital administrators. They could support both systems simultaneously with no sense of dissonance, since there was no real conflict between the technologies. Hospitals and community programs were perceived as serving different segments of the retarded population, and the hospital superintendents were glad of the opportunity to relieve themselves of chronic overcrowding and the constant pressure of the waiting lists.

In sum, the original innovations of the early and middle 1960s cannot reasonably be interpreted as a radical change in fundamental technology. Little more than a change in funding patterns was taking place, within the boundaries of well-established forms of service. Even this change was less dramatic than is supposed by the standard account: since the most optimistic estimates of the time placed the number of likely candidates for deinstitutionalization at around fifty to sixty percent of the combined hospital residents/waiting list count, which totalled less than three percent of the state's retarded citizens, the proposals were for a shift in distribution of retarded between hospital and community of only about one-and-one-half percent. This first step in the

change process is best understood as a conservative effort to maintain an established and valued system through correcting a flaw in its performance programs which had led to constant strain on the resources of the system and thus to consumer dissatisfaction. For this reason, there was virtual unanimity on the value of the innovations.

Nevertheless, some fifteen years later the strong anti-institutionalism depicted by the standard account had become a clear and major element of the system. Though no consensus had yet emerged, a significant segment of the attentive actors were loudly proclaiming that no amount of correction or improvement could salvage a system that was flawed in its most fundamental conceptions. The only solution, they argued, was the complete elimination of all the state hospitals, and the provision of all services through community-based programs.

If hospital- and community-based programs were not perceived as mutually exclusive alternatives in 1965, why should they be so viewed by many system participants in 1981? What happened during that period to cause such erosion of the power of the hospital technology to insulate itself from negative information? The following sections give a brief outline of some important processes that have progressively undermined the perceived viability of the hospital technology. As suggested earlier, the basic process seems to have been a feedback relationship between actions taken to correct flaws in peripheral features of the system and a rising sense of the ambiguity of the technology. As the perception of ambiguity strengthened, the organizational mechanisms which buffered the system against the impact of negative information weakened, resulting in a constant acceleration of the rate of criticism and the degree to which the focus of the criticism approached the core of the technology. The layers of the onion were being peeled away.

Professional Authority and Hierarchical Decision Structures

Earlier I described the California mental retardation system in 1960 as approaching March and Olsen's ideal type of a hierarchical decision structure. That structure is a powerful means of regularizing the bit of life that falls under the control of the organization, since it combines professionalism with a strongly-centralized command structure. In a pure hierarchical system members of a single profession control all of the important decision nodes of the organization. Since these persons have shared years of intense training in their profession, they are likely to share a strong consensus on the values, goals, methods and explanatory framework of the profession's technology. This training includes definition of what constitutes information relevant to professional decisionmaking, and the proper premises for evaluating its significance.

Any complex organizational system includes a large number of persons other than the core professionals who might claim a role in decisionmaking, including legislators, consumers, lower-level functionaries and members of the attentive public. In a stable hierarchical decision structure they do not make such a claim, largely because of the esoteric nature of the core technology. The assumption that those long years of training really are necessary to understand the technology fully and to employ it appropriately legitimates the exclusion of non-specialists from decisionmaking. Under these conditions, negative information concerning the core technology is unlikely to have an impact on the organization. The socialization of the professionals will prevent them from perceiving it as information at all, while non-specialists who are less insulated from heresy have no role in decisionmaking, are inattentive, and are usually willing to defer to the judgement of the professionals.

As powerful a process of regularization as this combination of professional socialization and hierarchical authority may be, it is extremely vulnerable to

the appearance of technological ambiguity. The authority of the professionals is dependent upon wide consensus on the value of the core technology, of which they are the privileged carriers. When the value of the technology becomes ambiguous, then there is reason to doubt that the professionals are any more qualified to make decisions than any other members of the organization. The legitimacy of excluding non-specialists with a real stake in the system from the decision processes falls into question.

Under such conditions, in a broader social system which places great value on egalitarianism it is likely that pressures will form to shift the decision structure from a hierarchical form to an unsegmented form. In that case, everyone will have the right to participate equally in every decision, not because the technology is so simple that all share competence, but because it is so ambiguous that no one individual's opinion is clearly better than anyone else's.

Of course, the essential question is why ambiguity should ever occur. I have argued that the immediate appearance of ambiguity concerning the fundamental principles of the technology is extremely unlikely. In the present case it is evident that the initial ambiguity was admissible because it concerned only a peripheral feature of the technology, the feeder mechanism which was perceived to result in inappropriate hospitalizations. The flaw was located in implementation of the basic principles, rather than in the principles themselves. Since even in a highly stable technology learning and change in peripheral features is a constant process, the innovations of the 1960s had no obvious revolutionary import for the system. However, the solutions adopted had the unanticipated consequence of raising an additional level of ambiguity, one that was a little bit closer to the core of the technology.

The Problem with Diagnosis

The process of planning and developing state-supported community services immediately resulted in two new sources of ambiguous information. As noted earlier, it had always been a basic tenet of the technology that only a small minority of the retarded required hospitalization. The inappropriate hospitalizations that were acknowledged to exist were at first considered to have resulted only from the aberration in the financial incentive structure of the feeder system, rather than in the more central technological issues of diagnosis of need and prescription of services. However, the efforts of the Department of Mental Hygiene in 1963 to ascertain the exact numbers of hospital patients and waiting list cases appropriate for alternative forms of service focused attention on how poorly the technology specified methods for diagnosing whether or not hospitalization was needed. The Department survey was conducted, under some very general guidelines, independently by the staff of each hospital. Their allocations of patients to the various types of services varied so wildly among the hospitals that it seemed impossible that the populations could really be that different. The authors of the report were forced to acknowledge the "subjectivity" of the methods (DMH 1963b:20).

This discovery appears to have caused no more than mild consternation, but it set in motion an escalating process. Within a year the Department had issued a new report which--though it apparently used exactly the same data--revised the percentage of patients requiring hospitalization upward from 37 to 64 percent (DMH 1964). The Assembly Subcommittee noted this discrepancy with obvious annoyance, and seemed reluctant to accept the validity of the new estimates (ASMHS 1965:48-51). Cracks had begun to appear in the credibility of important features of the technology among the gatekeepers to a necessary resource.

Tracing the subsequent history of efforts to improve the diagnosis of needs for services is far too complex an issue for the present forum. In gross outline, the effort has focused on the development of standardized testing instruments, the primary impact of which has been to expand and refine the perceived ambiguity of the technology. Effective instruments have been developed, but the appearance of each has set off a new round of arguments over the importance of the characteristics they measure. Today standardized tests are among the most controversial issues in the system, but the source of the controversy is less the quality of the instruments themselves than the ambiguity they reveal over what needs to be known and how such information should relate to service decisions.

Confusion over diagnosis has not resulted only from direct efforts to construct valid diagnostic instruments. The early modest efforts at moving mildly retarded hospital patients into the community also had the unanticipated consequence of heightening ambiguity over who needed hospitalization. George Tarjan, who was superintendent of the Pacific State Hospital during the early period, relates how the rapid discharge of the least disabled patients brought on an institutional crisis. Because of the capabilities of these persons the hospitals for decades had been able to rely on them for free provision of menial labor such a janitorial services and personal care for the physically disabled patients. When most were released in a short period, the hospital faced a labor crisis which was unlikely to be resolved by increased legislative appropriations for regular staff.

Out of necessity the hospital was forced to turn to its more severely retarded patients, particularly to those suffering from Downs' syndrome (then known as "mongolism"). They did not do so with much hope: the technology specified that these people were incapable of much beyond the most rudimentary development, an assumption confirmed daily by their behavior in the hospital.

However, their first efforts to train Downs' syndrome patients to feed patients suffering from severe cerebral palsy were so successful that they expanded the experiment and rapidly discovered that these patients had potential far beyond what they had expected, and in many cases were viable candidates for community programs (Tarjan, personal communication 1981). The dividing line between hospital and community cases had become even more confused. Today the system is divided between camps that say, on the one hand, that hospitalization remains necessary for some patients, and, on the other hand, that all patients should be in community facilities; but even within the first camp, few would profess to know quite where to draw the line.

Choice and Ambiguity

The second early source of the rising sense of ambiguity was the proliferation of choices among competing technologies. A stable closed system limits the choices available in an organization, in part by specifying a narrow range of activities that may be perceived as effective and appropriate for achievement of the official goals. The convergence of professional training and hierarchical power structures provides the mechanisms necessary to enforce conformity with this definition of the proper work of the organization. But when certainty begins to erode, and the authority of the core professionals is brought into question, this control mechanism is weakened. Competing technologies may begin to proliferate, often sponsored by professional groups formerly not represented (or at least not powerful) in the system, which are now vying for centrality in a redefined core technology. In the view of the consumers and the wider attentive public the mere existence of these alternative technologies increases the sense of ambiguity, because it requires choices to be made where none was necessary before. No agreed-upon rule for choosing among them is evident, particularly

if the competing professionals become willing to offer public criticism of one another.

This expansion of choices started modestly in the California mental retardation system, as would be expected where professional control of the system is strong. Community services in one form or another had existed for a long while, but were chronically underfunded (in the case of programs run by parents' organizations) or dominated by the hospitals (in the case of aftercare services for discharged patients). They had never been particularly attractive to professionals, and no strong group support had developed for those who did staff them.

The creation of the regional center system and associated events, however, stimulated a frenzy of professional organizing. The flow of state and federal funds into new forms of service not dominated by the psychiatrists who controlled the hospitals afforded opportunities for less powerful professions to stake out claims. Pediatricians, psychologists, social workers, physical therapists and others paraded before the hearings of the Study Commission and the Assembly Subcommittee to tout the role they should play in community services. Discovering that over 77 percent of the hospital patients were listed in the statistical reports of the Department of Mental Hygiene as retarded due to unknown causes, they argued that retardation was such a complex phenomenon that no one profession could possibly encompass all of the diverse research and forms of service needed. In its final report the Study Commission advocated involvement by many professions and the need for "inter-disciplinary teamwork and respect" (1965:94).

In almost all cases these other professionals were careful to state their roles as being supplements to the hospital technology of the psychiatrists, rather than as replacements, intended for the benefit of those retarded who would not be hospital patients. Thus, they avoided any direct challenge to the

to the established technology and its carriers, while focusing their efforts on building constituencies for their services among consumers and parents. However, over time, as the strength of the community services grew these other professionals became among the most vocal critics of the traditional hospital system.

Today the system is marked by a multitude of competing services and a chronic inability to assess their value or to match service type to diagnosis systematically. The administering agency has attempted to define which services should be considered "basic" and which "frills", but these efforts have met with outraged resistance and a flat denial of the validity of the distinction.

Growth in the Base of Participation

The processes described in the preceding two sections had the effect of greatly expanding the number and diversity of people participating in decisionmaking in the system. This expansion had added its own increment to the growth of ambiguity. Many of these persons had not been previously socialized into the closed system of the hospitals. They came from diverse backgrounds, and in the case of most parents had no professional training in mental retardation at all. Thus, the common ground of discourse was eroding. Meetings of organizations in the system are now frequently dominated by argument and controversy in which progress is seldom made.

The Equity Principle

There are three basic strategies for allocating resources in a service system: (a) service according to need; (b) service according to demand; and (c) service according to equity (i.e., distribute resources as equally as possible among the eligibles). Strategies in actual cases are usually mixtures of these

ideal types, but differentiating them in the pure form is useful in illuminating certain of the processes operating in the case study.

These three strategies may be ranked in order of the cognitive abilities they impute to the core professionals of the organization. The greatest cognitive demands are implied by the need strategy. This strategy inevitably results in an unequal distribution of resources among clients, in response to the differing diagnoses of their extent of need. In order to obtain general agreement on the justice of this unequal distribution of public goods, it is necessary that there be consensus among the relevant actors on a technology for (a) diagnosis of need; (b) translation of the diagnosis into a prescription for treatment; (c) prioritization of needs under conditions of scarcity, and (d) accreditation of professionals to implement the technology. It requires considerable faith in the competence and stability of the organization and its technology, and thus is the service strategy most consistent with a high degree of professionalization. The need strategy, then, may be found most frequently associated with hierarchical decision structures.

The cognitive requirements of the demand strategy are less extensive. The diagnosis of need is largely externalized by assuming that those who appear demanding services are those in greatest need, or whose needs most accurately fit the services the organization is prepared to offer. Where the need strategy may be seen as closely related to professionalism, operating in an environment in which service technologies and deployment of resources are seen as non-problematic and uncontroversial, the demand strategy is related to politics, in an environment marked by faith in the value of the core technology but strong controversy over the distribution of services. In the face of militant demands for service, the demand strategy aims at keeping the political peace by greasing the squeakiest wheels. Under a demand strategy, professionals may continue to

operate the routine core technology, but politicians controlling program funds or politically-sensitive professional managers exercise considerable control over the access point to the process, defining eligibility criteria that are not strictly related to the internal logic of the core technology. The tendency is thus for demand strategies to be found in association with specialized decision structures.

The equity strategy implies the most modest cognitive demands on the agency, since it limits the role of the agency largely to certifying applicants as members of the pool of eligible recipients who once certified are able to assert a "right" to a roughly equal share of the pot. In a sense the equity strategy is as political as the demand strategy, but the political situation is significantly different: in the demand strategy certain groups or individuals are particularly successful and command a disproportionate share of the resources, while in the equity strategy the competing groups and individuals are sufficiently balanced in power that no one can command a share greater than its proportion of the total population.

Where a need strategy implies strong technological certainty to justify the distribution of resources on the basis of professional diagnosis, the equity strategy seems particularly consistent with conditions of technological ambiguity. It is extremely difficult under such conditions to get agreement on important decisions--and none is more important than who gets what benefits--simply because there are so many alternatives to choose from, so little basis for telling which one is going to solve your problems, and so little confidence in the advice of experts, who in any case will appear in some numbers lined up behind each one of the possibilities. Also, the problem of getting agreement is made even more difficult simply by the increased number of people participating in the decision process, each one of whom has to be at least partially satisfied. These conditions

are the same which push the decision structure toward the unsegmented form, with its emphasis on egalitarian participation in decisionmaking.

There is a decision rule that is capable of attracting wide consensus under conditions of ambiguity, and that is the rule of equity. Since any group's chance of commanding the lion's share of the resources is very slim, all have an interest in at least ensuring that each will receive an equal slice of the pie.

The emergence of this kind of egalitarian notion, expressed both in the development of the unsegmented decision structure and the principle of equality in the distribution of benefits, is attributed in the standard account to the effects of the civil rights movement. I suspect that it is instead simply a temporary aspect of the change process itself--another index of rising technological ambiguity--which may be expected to disappear (or at least to recede from centrality) onee a new, reconceptualized system has gained widespread support. Certainly it is true that egalitarianism is an important value of American society, but this does not mean that the changes in the decision structures and rules of social service bureaucracies were caused by a resurgence of populism in the wake of the civil rights movement. There is also a strong value in American ideology on utilitarianism, and so long as there is a stable technology in existence inequality in either access to decision processes or the distribution of resources is not considered problematic. The accepted technology certifies the superior value of professional opinion and gives reasons for diagnosing clients as having widely different claims upon the resources of the system. But once ambiguity enters the system, the ambient value on egalitarianism may rush in to fill the void left by a disintegrating technological consensus. When a new technology achieves general acceptance, the conditions favoring the utilitarian ethic will reemerge, and decision structures will probably return to a segmented form, while unequal distributions

of resources will once again be taken as sensible expressions of professional diagnosis of need.

A final symptom of the spiraling ambiguity in the system is the stress placed upon legalistic advocacy for "clients' rights". In an unsegmented decision structure, where diverse interests are incompetition and overarching principles (other than equity) for deciding among rival claims for resources are absent, adversarial relationships are an almost inevitable mode of interaction. Since professional diagnosis of need has been largely discredited, clients or their representatives who are attempting to determine whether the client is receiving proper attention must turn to a process of "social comparison" (Festinger 1954). That is, they are less concerned with the relationship between formal diagnosis and services rendered than with comparing the services rendered to the focal client with those given other clients. Thus, advocacy agencies publish lists of services available and admonish clients to make certain that they receive their share, and demand standardization of service availability across agency locations. The question asked is not, did the client receive what he needed? but did the client receive what he deserved?--his fair share of the resources of the system.

In this light, it is clear that the identification of these services as "rights" rather than "needs" of the clients is another linkage in the feedback loop driving the rise of ambiguity. Under the old stable closed system professional diagnosis of need was recognized as the only reasonable basis for the prescription of specific services. The growing sense of ambiguity undermined the credibility of such diagnosis, giving rise to the principle of equity and the stress on advocacy. In turn, these features of the system reinforced anti-professionalism. The old system sanctioned inequalities in the distribution of resources. As that system was increasingly discredited, the advocates came to suspect that it was never more than a sham, disguising the personal or class

interests of the professionals which were now assumed to have been the real motivation for unequal distribution of resources. Initial doubts concerning the efficacy of the professionals' methods grew into outright hostility and deep suspicion of their motives.

But with anti-professionalism so complete, what is the basis for claiming that client X should receive service Y? In the absence of a stable technology, the only argument available is that of rights based upon social comparisons: present clients have a right to Y services because previous clients of the system received Y services, or because members of the general population receive Y services. (In the latter case the argument is actually reversed: previous clients did not receive the service because professionals certified that they did not need the services received by "normals", or could not benefit from them. Since the professionals' opinion is no longer trusted, then the services should be rendered.)

Conclusions

Under what circumstances, and through what processes, does the acquisition of information contribute to fundamental organizational change? I have tried to show that fundamental change presents a special problem in analysis of the use of information in organizations. Because such change involves a shift in the deep structures which organize the way in which members recognize (or ignore), process, and interpret information, use of information in fundamental change is a qualitatively different issue from that of information use conducted routinely under the guidance of a closed system.

Because of the powerful information-organizing capacity of the closed system, it is unlikely that change could begin with the acquisition of information that is perceived by members of the organization as challenging the fundamental

assumptions of the system. In this case study information that questioned whether hospitals did not hurt rather than help the retarded was received early in the change process, but was easily interpreted by almost all involved as information which supported long-established principles of the closed system. Thus, it was treated as a minor revelation dealing with peripheral aspects of the technology.

At the same time, a great deal of attention was paid to information which questioned the viability of continued expansion of the hospital system. This information, relating to the costs of maintaining that program, and to the problematic nature of the financial incentive system which appeared to drive the program, was interpreted by the actors in a way that seems quite accurate to the observer years later. But these programs lay near the periphery of the system, and so were not strongly buffered against scrutiny and change.

A decade after the change process had begun, ambiguity had penetrated deeply into the closed system, and the information on the damaging effects of hospitalization--which were easily interpreted away in the early 1960s--had come to be perceived by many as proving that hospitals were harmful to <u>all</u> retarded, and should be completely eradicated. In the early 1970s Governor Ronald Reagan actually proposed complete elimination of the hospitals (though apparently for budgetary rather than humanitarian motives); however, the constituency for this point of view was still outnumbered by those who continued to interpret the information as referring only to the less severely disabled, and the proposed shutdown was rescinded. Today opinion reamins deeply divided, but the general impression is that anti-institutionalism continues to gain ground.

I have argued that receptivity to information challenging the fundamentals of a closed system is unlikely to be the initial cause of a process of fundamental change; rather, it is something achieved as a late product of a long

and painful transitional process, and shortly precedes a radical reconceptualization of the field.

A major problem seems to remain. Since no organization functions perfectly, there is always negative information about the peripheral performance programs of every organization, and incremental change reacting to this information is a more-or-less constant feature of almost all organizations. Why, then, did this particular set of changes at the periphery set off the process of spiraling ambiguity that has been dismantling the closed system, when a multitude of other incremental changes over the years of stability did not? Perhaps there are just certain weak points that lead naturally into a transitional period whenever change strikes them. However, I do not find this a very satisfying explanation.

I suspect that the answer lies in the relationship between the organizational system and its environment. Many of the fundamental assumptions of any publicly supported or tolerated institution deal with general theories of humans, society or nature that are shared with the wider public. So long as that environment remains stable in its appreciation of those assumptions, then it is likely that peripheral changes which touch briefly upon deeper assumptions will have no major impact upon the stability of the closed system. But if public opinion abandons the closed system, then routine adjustments may suddenly produce more dramatic unanticipated consequences when they force actors to recognize that their reality has become outmoded.

In the case of mental retardation services, it is likely that the general context of the civil rights movement played this role, by undercutting the assumption that it was proper for society to segregate people who are "different". At most times a shift of one or two percent of the retarded from institutions to the community would have had little effect on the perception of the system; but here it drew attention to the fact that people had been held in institutions

for years who were now seen to be quite capable of life in the community. This led to a closer examination of the premises of the system which continues today.

This would suggest that fundamental organizational change is likely only in times of ferment in the broader society. A similar pattern was found in my study of fundamental change in American prisons (Rounds 1979), but only additional case studies can determine how general that pattern might be.

NOTES

1. This should not be taken as a theory of how the onion is originally created. The paper is about how closed systems die, rather than how they are born. In general it may be said that the whole onion--or at least a crude version of it--is formed all at once, but then afterwards growth occurs almost exclusively at the periphery. But of course the matter is far more complex than this, and cannot be dealt with adequately here.

2. The articles on deinstitutionalization in mental retardation services cited below are, for the most part, focused on the process throughout the United States, rather than specifically on California. I think it does not misrepresent their positions to take them as intended to reflect the California case accurately. However, it should be noted that deinstitutionalization was occurring simultaneously in most of the states, and treating the California case in isolation necessarily oversimplifies the change process. This account makes no claim to be a complete explanation of that process.

3. This apparent duplication of effort was the result of a long-term political rivalry between Governor Brown and the Speaker of the Assembly.

4. I use the three percent figure in these calculations because that was the figure commonly accepted in the 1960s. However, more recently a lower figure of around one percent has been proposed (Tarjan, et. al. 1973).

REFERENCES CITED

Assembly Subcommittee on Mental Health Services

1964a A Preliminary Proposal to Eliminate "Waiting Lists" for State Hospitals for the Mentally Retarded.

1964b Supplementary Factual Report #1

1964c Supplementary Factual Report #2. A Survey of Families with Children on the Waiting List.

1965 A Redefinition of State Responsibility for California's Mentally Retarded.

Begab, Michael J.

1963 Community Programs for the Mentally Retarded and Their Families. In Coordination of Services for the Mentally Retarded. Barbara Hunter, ed. Pp. 24-42. Sacramento: Department of Mental Hygiene

Bourdieu, Pierre

1977 Outline of a Theory of Practice. Richard Nice, trans. Cambridge: Cambridge University Press.

Cain, Leo F. and Samuel Levine

1961 A Study of the Effects of Community and Institutional School Classes for Trainable Mentally Retarded Children. San Francisco: San Francisco State College.

Davies, Stanley P.

1930 Social Control of the Mentally Deficient. New York: Crowell.

Department of Mental Hygiene

1963a Hospital Bed Needs for Mentally Retarded Patients in California, 1963-1968.

1963b Survey of Patient Needs for Residential Care and Assistance: Hospitals for the Mentally Retarded [and] Preadmission Services.

1964 Tentative Proposal for a Five-Year Program.

Festinger, Leon

1954 Theory of Social Comparison Processes. Human Relations 7:117-140.

Garner, Roberta

1977 Social Movements in America. 2nd Edition. Chicago: Rand McNally.

Halpern, Joseph, R. Binner, C.B. Mohr and K.L. Sackett

1978 The Illusion of Deinstitutionalization. Denver: Denver Research Institute.

Kuhn, Thomas S.

1970 The Structure of Scientific Revolutions. 2nd Edition. Chicago: University of Chicago Press.

1977 The Essential Tension: Selected Studies in Scientific Tradition and Change. Chicago: University of Chicago Press.

March, James G. and Johan P. Olsen

1976 Ambiguity and Choice in Organizations. Bergen: Universitetsforlaget.

March, James G. and Herbert A. Simon

1958 Organizations. New York: John Wiley and Sons.

Moore, Sally F. and Barbara G. Myerhoff, eds.

1975 Symbol and Politics in Communal Ideology: Cases and Questions. Ithaca: Cornell University Press.

Rounds, J.

1979 Social Theory, Public Policy and Social Order. Ph.D. dissertation, University of California, Los Angeles.

Simon, Herbert A.

1976 Administrative Behavior. 3rd Edition. New York: Free Press.

Sproull, Lee

1981 Beliefs in Organizations. In Handbook of Organizational Design. Paul C. Nystrom and William h. Starbuck, eds. Pp. 203-224. London: Oxford University Press.

Study Commission on Mental Retardation

1965 The Undeveloped Resource: A Plan for the Mentally Retarded in California. Sacramento: State Printing Office.

Tarjan, George, Richard K. Eyman and Harvey F. Dingman

1966 Changes in the Patient Population of a Hospital for the Mentally Retarded. American Journal on Mental Deficiency 70:529-541.

Tarjan, G., S.W. Wright, R.K. Eyman and C.V. Keeran

1973 Natural History of Mental Retardation: Some Aspects of Epidemiology. American Journal on Mental Deficiency 77:369-379.

Taylor, Steven J. and Robert Bogdan

1980 Defending Illusions: The Institution's Struggle for Survival. Human Organization 39:209-218.

Ziman, John

1978 Reliable Knowledge: An Exploration of the Grounds for Belief in Science. Cambridge: Cambridge University Press.

The Nature of Managerial Attention

Lee S. Sproull

Carnegie-Mellon University

Prepared for the Carnegie-Mellon University Symposium on Information Processing in Organizations, October, 1981. The writing of this paper was supported in part by grant G-79-0018 from the National Institute of Education and by many conversations with good friends and helpful colleagues. I would like to acknowledge particularly Jim March and Bob Sproull.

Managerial information processing has been a target of systematic empirical investigations for nearly thirty years. These investigations have focused primarily on how particular kinds of important, usually decision-relevant, information are processed, for example, information about budgeting or strategic planning.[1] By abstracting these information-processing behaviors from the daily organizational context within which they occur, these investigations, obscure how managers process the unending stream of information--important and unimportant on a wide range of topics--that managers encounter every day. To understand managerial information processing more generally, it is necessary to focus on the manager as a continuous processor in his information environment. This paper explores managerial attention as a way to understand the nature of the continuous interaction between the manager and his information environment.

The paper is organized in three sections. The first discusses managerial attention within the context of information processing. It points out some difficulties with investigating attention in natural settings and discusses the strengths and weaknesses of previous relevant research on managerial time allocation. The second section describes an empirical study of managerial attention. The third section discusses some implications of this research for how we think about information processing in organizations.

MANAGERIAL ATTENTION

Attention is defined as allocating information-processing capacity (receiving, cognitive processing, disseminating) to environmental stimuli over time.[2] When a manager is attending to a particular stimulus he is processing information relevant to it. Leaving aside yogis, sleepwalkers, and others in

trancelike states, managers are always attending to something. Note that this definition is broader than the psychological concept of attention. It includes the overt, visible, steps of receiving and disseminating information that are captured in the common-place phrase, "paying attention to," as well as the mental steps of noticing and encoding. Note also that it is possible to investigate managerial attention as it is defined here without understanding explicitly the internal cognitive processes of interest to experimental psychologists.[3] That is, one can study how, to whom, and to what managers devote their attention throughout the day without considering how those stimuli are mentally coded or represented. Although the internal, mental aspects of attention remain hidden, the concept can prove useful in systematic investigations of managerial information processing.

Attention has both structure and content. At any moment the content of managerial attention is that which the manager is processing: the substance or particular kind of information--for example, information about sales or supplies. The investigations mentioned at the beginning of this paper focus on the content of attention. The structure of managerial attention is how the manager processes: through what modality, in what location, for how long--for example in a telephone call with a lawyer or on a scheduled plant tour with the head of operations. Certain kinds of information (content) are more likely to be conveyed in certain forms (structure)--for example, detailed financial information in written reports prepared by certain people or political gossip in private conversations with other people. The interaction between structure and content is likely to affect the cognitive behavior of the manager. For example, a manager may be attending to financial information (content) but whether he is

doing so through reading an investment prospectus or hearing an exultant phone call from the head of sales (structure) influences whether he engages in detailed numerical calculations or simple aspiration-level comparisons.

Investigating managerial attention has the potential to improve our understanding of information processing in several ways. It can illuminate something of the context within which decision-relevant information processing occurs. That is, we can ascertain what fraction of total managerial attention is devoted to this kind of processing. More interestingly, we can learn how attention to decision-relevant information is distributed throughout the day. It can provide a useful characterization of the actual information environment within which managers operate. And it can help those who want to improve or assist managerial information processing by giving them some sense of relatively frequent or infrequent patterns of attention. One common complaint against management information systems, for example, is that they are designed to help managers do something that managers have relatively little interest in doing and pay little attention to (Ackoff, 1967), namely search systematically through large amounts of archival information. Some insight into managerial attention could lead designers to focus on the kinds of information-processing activities that managers engage in frequently.

There are complications with attempting to understand managerial attention in a natural setting. Because we have no direct access to the manager's mental processes, attention must be assessed through visible and audible cues. One of the most common cues to attention is physical presence, but presence and attention are not synonymous. A General Motors group vice-president illustrates

the distinction in his description of executive committee meetings (Wright, 1979: 27-28). All participants were required to read before each meeting verbatim copies of any presentations to be made at the meeting; thus they could, and did, sleep through the actual meeting. They were present, but they were not paying attention. Researchers who rely on calendar data can be particularly troubled by this complication. Although they can know where and with whom managers spend their scheduled time (structure), they cannot know to what they are attending or even if they are awake (content). This complication can be lessened by collecting multiple cues to attention, for example calendars, self reports, and interviews with others.

A second complication arises because there is no sound metric of attention. The most common empirical metric--elapsed time--can be misleading. All minutes of attention need not have equal weight; managers can exert more information-processing effort ("pay closer attention," in lay terms) at some times than at others. Professors commonly observe that undergraduates pay closer attention in class on the day before an exam than they do on other days; managers no doubt similarly pay closer attention to certain people, topics, situations, than others.[4] It is difficult to reduce this complication in a natural setting. One possible solution would entail weighting time allocation data after the fact by an effort or importance factor; the prospect of such a weighting exercise should chill the warmest empirical heart. Fortunately, there is more than enough to be learned by invoking the assumption that all minutes of attention have equal information processing weight.

A third complication is that attention can be leveraged; a manager can

direct his agents to attend to certain issues then report back to him. In this way managerial attention can produce an information-processing multiplier effect. Easing this complication simply requires that researchers be explicit about their focus of attention, direct or leveraged attention. Again, currently there is more than enough to be learned by taking the simpler path.

The organizational researchers who have come closest to empirical investigations of managerial attention are those who study time allocation.[5] Through diaries, checklists, interviews, or observation, these researchers estimate the amount of time managers devote to various activities during the work day. The most positive contribution of these studies is to challenge with relatively systematic data on actual time expenditures an idealized view of managing. The major shortcomings of these studies stem from their preconceived definition of managing as a particular kind of behavior managers engage in. These studies focus on the "job" or "work" of managing, and ignore the rest of the day's activities.[6] Focussing on "job" or "work behaviors" rather than on all managerial activity entails three difficulties. The first is that the empirical distinction between "job-related" and "non-job-related" behaviors is arbitrary and may be difficult to draw. Second, ignoring certain behaviors leads to underestimating the manager's total time budget during the workday. Third, focussing on "job-related behaviors" rather than on all behaviors ignores the continuity of time--the fact that "there are no behavioral vacuums in the life of an individual" (Atkinson, 1969: 105)--and makes the manager look rather more like a puppet than like a ballet dancer.[7]

Because these studies focus on the job rather than on the manager as a

continuous information processor, they move fairly rapidly to general statements of managerial "purpose," "programs," and "roles." Even those who collect finely-grained data engage in what one of them acknowledges as "a somewhat arbitrary partitioning of the manager's activities into affinity groups" (Mintzberg, 1973: 55). The affinity groups of activities are then assigned such general labels as "monitoring the environment" or "handling disturbances" or "planning." The difficulty with such general labels is that most managerial activity can be labeled by more than one purpose category. Hence, assigning single labels is misleading. The most confusing instance of such labeling practices is using the label "information processing" to categorize some behaviors but not others, thereby implying that when managers are "leading" or "negotiating" that they are <u>not</u> processing information.[8] These shortcomings in previous time allocation research can be overcome by focussing on the manager as a continuous information processor (rather than on his job of managing), using a relatively fine-grained level of detail, and avoiding the temptation to assign general-purpose categories to the managers' behavior.

Because there is inevitably more information in the manager's environment than there is processing capacity for it, how the manager allocates his attention among competing demands determines what information gets processed and what gets ignored.[9] But allocation rules alone would be insufficient to characterize attention. We also need to understand how allocation is affected by interruptions. The connection between allocation and interruption of information processing suggests an analogy with a computer job scheduler. The scheduler is responsible for allocating processor time among runnable programs so as to maximize an objective function, for example the amount of useful work

done or computer center revenues, subject to both capacity and external constraints. But allocation decisions can be affected by input/output interrupts. Even as the scheduler allocates computer time to programs, the operating system can interrupt a running program to do input/output. Each operating system has a hierarchy of interrupt priorities; thus interruptions themselves can be interrupted. In practice schedulers do not solve the optimization problem for two reasons: The calculations would be too expensive and there is insufficient information about the future behavior of programs. Each scheduler is simply tuned to work acceptably well through ad hoc techniques and parameters. The analogy suggests that managers may well operate under predictable, though perhaps unarticulated, interrupt allocation rules.

This discussion suggests three general empirical questions with implications for managerial information processing: What are the structure and content of managerial attention? What are the allocation rules? What are the interrupt rules? The work reported below addresses these questions through a strategy of observing managers and recording data about the structure and content of their attention on a minute-by-minute basis. This strategy yields an answer to the first question directly, for one set of managers. Although not strictly comparable, the data can be set in the context of previous time allocation data to address the first question more generally as well. The second and third questions are addressed through inferring rules from the observation data.

EMPIRICAL WORK

Sample and Methods

Seven first-line managers of new programs in public sector organizations were selected as subjects through soliciting expert nominations of important programs and successful managers. Each manager was told that the researcher was interested in how new programs were implemented and wanted to observe the managers as they went about their daily activities. All managers agreed to participate. (Table 1 presents descriptive statistics about the managers, their programs, and their larger organizations.) Each program was staffed primarily by professionals--teachers or lawyers. Program staff size ranged from six to sixteen; annual program budgets ranged from $160,000 to $900,000. The larger organizations housing the new programs were four school districts, one community college district, and a state agency responsible for labor relations in education. Two of the managers were located in the state agency; each other organization contributed one manager.

The managers were observed for a total of twenty-nine working days (between 3 and 6 days per manager), during which time descriptions of all of their activities one minute or longer in duration were recorded.[10] Each description included data on both the structure (where, when, how, with whom, how initiated) and content (subject matter) of attention.[11] The descriptions were written in real time and transcribed each evening. Each day's observation record was entered in a computer text file to be analyzed through a series of text-processing programs.[12]

Findings

The structure of attention can be characterized by four dimensions: length and location of workday, number and duration of attention episodes, medium of attention, and interaction and initiation. The average workday is about 8.7 hours long.[13] About 74% of the work day is spent in the program offices; 15% outside the program offices, usually in meetings with the manager's superior or other members of the larger organization; and 11% at lunch. The day is composed of many activities of brief duration: the average day contains 58 activities (average duration = 9.0 minutes). The average day includes 3 scheduled meetings (average duration = 60m.), 28 unscheduled conversations (average duration = 6.5m.), 14 phone calls (average duration = 4.8m.), and 13 periods alone (average duration = 7m.).

The manager spends about 80% of his day talking to people. (Table 2 indicates the proportion of conversation by medium and other participant.) Managers talk primarily to their program participants, usually in unscheduled conversations; second arily to members of the larger bureaucracy, usually in scheduled meetings; and least to peers and the public, usually in telephone calls. Managers initiate about half of these conversations: 51% of face-to-face conversations; 54% of the phone calls. During the 20% of the day that the manager is not talking to people, he is alone. These solitary periods are not initiated by a single individual: if the manager is alone, it is both because the manager chooses not to be talking with anyone at the moment and because no one chooses to talk with the manager. (How solitary time is terminated is described below under interrupt rules.)

To summarize the structure of managerial attention: it is local, choppy, mostly unscheduled, oral, and as much other-directed as self-directed. When we turn to other time allocation studies, we discover marked similarities. (Table 3 displays major comparisons.) The major differences seem to be associated with level of the hierarchy at which the manager is operating, with more time spent in longer and more meetings as hierarchical position increases.

The content or subject matter of managerial attention is distributed across four broad information categories: program, logistics, external environment, and interpersonal relations. (The remaining information is personal affairs and talking to the researcher.)[14] The distribution of attention to various kinds of information differs substantially between the state agency managers and the school program managers. (Table 4 compares the content of attention for the two kinds of managers.) The differences can be characterized generally by observing that agency managers attend more to the internal details of running the program while the school managers attend more to outside requests and social pleasantries. For example, the agency-based manager spends a greater proportion of his day attending to the program core (30% versus 12%). The agency-based manager also spends a greater proportion of his day attending to materials and supplies (23% of the day versus 6% of the day). The school-based manager spends a greater proportion of his/her day attending to external requirements and information (19% versus 4%). The school-based manager also devotes a greater proportion of his/her attention to social pleasantries (11% versus 4%).

One way to think about differences in the content of attention is in terms of differences in topics of conversation, since all managers spend about 80% of

their day in conversation. A new person enters the manager's field of attention on the average of once every 10 minutes throughout the workday; it is possible to infer allocation rules for the topic of their conversation in part from the content and structure data above. (Figure 1 illustrates how the allocation rules might operate.) About half the time, the other person initiates the topic; hence the first rule is, "If other person defines topic, manager attends to that topic." If the other person does not define the topic, managers usually dispense with old business or resume interrupted business first. (This includes agenda items or scheduled meetings.) Hence the second rule is, "If there is a previous topic to attend to, manager attends to update." Scheduling and coordinating then take precedence. The third rule is, "If coordination is necessary, manager attends to coordinating." If none of the above rules generates a topic for conversation, managers often will tell a story or ask, "How's it going?" This latter question of course then invites the other person to define the topic of conversation. Here the major difference between the state agency program and the school programs become apparent. State agency personnel almost always treat a "How's it going?" question as a request for performance information, which they then provide. School personnel, on the other hand, almost always treat that question as a social pleasantry, to which they respond with either a very general, "just fine," or a more specific report on their non-work lives.

The topic of conversations, as well as managerial attention while alone, both are subject to interruptions. In order to talk about interruptions it is helpful to define three units of attention. An <u>event</u> is any continuous period of time, one minute or longer in duration, in which neither the structure nor

content of attention changes. For example, a telephone call on a single subject constitutes an event. An activity is one or more events, which can be separated by intervening events, in which neither the structure nor content changes. For example a telephone call interrupted by someone asking a question constitutes two activities (telephone call and question), but three events (first part of phone call, question, second part of phone call). A topic is all activities in which the content of attention does not change. All meetings, conversations, phone calls, and desk work devoted to a single subject constitutes a topic. The average number of events during the manager's day is 73 (average duration = 7.1m); the average number of activities is 58 (average duration = 9m); the average number of topics is 24 (average duration = 22m.). The difference between the number of events and the number of activities in a day characterizes the day's interruption rate; on average 21% of the day's activities are interrupted. It is important to note that managers spontaneously interrupt themselves as often as they are interrupted by others.

Interruption activity can be summarized by the following rules:

- 1. Attend to what is in front of you, in descending order of interrupt priority:

  a. Ringing telephone
  b. Person entering field of attention (e.g., coming into office)
  c. Person currently in field of attention (e.g., sitting across the desk)
  d. Problem of previous phone call or visitor
  e. Papers on the desk

- 2. To change what is in front of you:

  a. If 8 minutes pass without someone's changing your focus of

attention, change it yourself
b. If 9 minutes of conversation pass without interruption or termination, interrupt or terminate it yourself.
c. If 15 minutes pass without interruption while you are alone, find someone to talk to.

Discussion

At least three features of managerial attention demonstrated in the above data are important for understanding information processing in organizations. The first is that brief, oral communication dominates all other information-processing behaviors. The second is that much of the information a manager processes during the day is either mundane, decision-irrelevant, or both. The third is that the content of managerial attention varies across different kinds of organizations at the same hierarchical level.[15]

Given these features, several questions emerge. How do managers notice and evaluate decision-relevant information? Two cues that the managers in this sample seemed to use were deadlines and carriers, certain trusted people. When a decision deadline was close at hand, managers looked and listened more attentively for relevant information. By the same token they paid less attention to potentially relevant information when deadlines were not near. Managers also seemed to use certain people in their environments as sources, seeking them out for advice or opinions. One interesting aspect of this sourcing behavior is that sources could, and did, deliver erroneous information, but because it was usually conveyed orally it was not easy to check for accuracy.

How do managers accumulate (notice and remember in a usefully-organized

fashion) disparate bits of information? When information is conveyed orally and at different times it can be difficult to notice patterns or trends. Often the same request would be handled eight or ten times by a manager as a special case before he would realize he had heard it before and could develop a procedure to handle it. None of the managers I observed made use of simple aggregating or enumerating procedures such as logging the topic of telephone requests.

How much control do managers have over their own allocation and interrupt rules? Sometimes they came in at 7 am to work for an hour or two before the interruptions began. Sometimes they told secretaries to hold all calls or not to let in any visitors, thereby temporarily overriding either the first or second interrupt rule. But do they ever change the default option on their rules or generate new rules?

This study was limited to only seven first line managers in only two kinds of organizations. Obviously its empirical limitations suggest caution in speculating on the significance of its findings. The findings on the structure and content of attention, however, are quite consonant with those of time allocation research in other kinds of organizations and at other hierarchical levels. What is needed now is more work on allocation and interrupt rules.

IMPLICATIONS AND SPECULATIONS

Interruptions and interrupting are a good candidate for future research. The work reported here suggests at least two cautions about the nature of interruptions, however. The first is related to the relationship between persistence and interruption. Common formulations and advice assume that

managers would persist in whatever they were doing if the environment did not generate interruptions.[16] This stance is predicated on an inertia model of attention. In contrast to the common assumption, in the study reported here managers generated their own interruptions--interrupted themselves--at least half of the time. In these instances, their attention seemed to decay over relatively short periods of time. Thus perhaps one should consider, as well as an inertia model, a stochastic model of attention.

The second caution is related to the meaning and measurement of interruptions. If an interruption is defined as a brief break in the main flow of activity, it is not possible to identify interruptions as they occur. This is the case, in part, because when the manager's day is composed of many brief disparate activities it is not clear what the "main flow" is. Additionally, if a break in an activity occurs, it is only after it is over and the manager returns to the previous activity that the break would be labelled as an interruption. It is for these reasons that the three units of event, activity, and topic were constructed for the research reported here.[17] The difficulty in distinguishing between interrupted and interrupting activity suggests that it may be misleading to think about managerial attention in terms of task and interruptions; it may be more sensible to think about the manager as a multi-task processor. The analogy with a computer still holds; it is often the case that the scheduler allocates run-time to a program over and over, swapping it in and out of core many times, before the program terminates. Similarly managers seem to keep several tasks going concurrently, swapping them in and out of their immediate focus of attention repeatedly throughout the day.

This conception of managerial attention suggests several research possibilities. Psychological research on problem solving behavior mostly focuses on solving a single problem, albeit sometimes with interference. Does concurrent multi-task problem solving look just like a linear combination of single-task problem solving? Probably not. What is the effect of swapping behaviors on start-up processes, memory, problem formulation, solution time, etc.? Organizational research on problem solving also usually focuses on solving a single problem, or one problem at a time.[18] In an organizational context, what mechanisms do managers use to keep from losing tasks? What effect do their allocation and interrupt mechanisms have on the likelihood of new information or important information reaching their attention?

The discussion of multi-task processing is predicated on an instrumental view of managerial behavior. Other views accommodate the allocation and interrupt behaviors reported in this study and should not be overlooked. Managers may be guided by the desire to establish and maintain attention obligations. In this view, most of the time the content of conversations is not important. What matters is that contacts are being trained to be accessible when they are needed. Or managerial attention may be influenced by rules about the presentation of self. One such rule that would explain why managers spend much of the day in short conversations is, "Stop talking before you make a fool of yourself."

NOTES

[1] There has been for example, research on how managers process information relevant to making resource allocation decisions (Cyert and March, 1963; Crecine, 1969; Larkey, 1979), to identifying new problems (Pounds, 1969), to understanding the business environment (Keegan, 1974), to predicting or explaining the behavior of enemies (Wohlstetter, 1973; Allison, 1971; Shlaim, 1976), and to constructing a clear mental picture of the organization (Simon, 1953; Clark, 1972).

[2] Simon (1971;1973).

[3] Attention has had an uneven career in psychology--popular in the nineteenth century, out of favor in the first half of the twentieth century, reviving in the 1960's and 70's with the emergence of cognitive psychology (Kahneman, 1973: 1-2). But unless one is interested in eye movements, attention remains a fairly slippery psychological concept, which, according to one observer, "invites uncertainty and chaos" (Mostofsky, 1970: 14-15).

[4] See Kahneman (1973) for an extensive review of the experimental literature on the relation between attention and effort. He also comments on the absence of a satisfactory metric for attention and suggests that physiological measures of effort may be developed (pp. 4-5).

[5] Economists have explored formally and through simulation how decision-makers might allocate attention to alternative decision-making opportunities (Radner, 1975; Radner and Rothchild, 1975; Freeland and Stabell, 1978). In these analyses behavioral questions such as how managers notice and evaluate

opportunities are, for the most part, ignored. So too are all decision-irrelevant situations. Some sociologists have investigated how people allocate attention among work and various leisure activities (Szalai, 1966; Szalai et al., 1972; Robinson, 1977). Work is treated as an undifferentiated activity however; they have not investigated allocations within the work component. See Mintzberg (1973) and McCall, Morrison and Hannan (1978) for detailed summaries of the managerial time allocation findings.

6 Researchers have studied "job behavior episodes" (Dubin and Spray, 1964); "work activities" (Horne and Lupton, 1965); "job behavior incidents". (Stewart, 1967). Even Mintzberg's (1973) generally thorough and important work omits all "contacts the manager had with his secretary" (271), "telephone calls from the manager's wife" (271), and "work that was ostensibly social in nature ([for example,] golf club board meeting)" (272). One exception is Cohen and March (1974) who analyzed all the recorded activities of American college presidents between the hours of 8am and 6pm.

7 Of course even ballet dancers look jerky if their movements are analyzed in very fine detail, as for example through stop-action photography. See Spence (1978) on the problems of trying to describe continuous behavior.

8 This kind of general categorizing also tends to produce subsequent research that focuses on the categories rather than on continuous behaviors and hence falls into unproductive debates over, for example, whether or not managers actually "plan" or the relative frequency with which they "monitor" as opposed to "make decisions." See Snyder and Glueck (1980) for an acute example.

[9] Economists suggest that a fire-fighting rule (allocate effort to the worst problem in a particular time period) produces the best performance over time (Radner, 1975; Radner and Rothchild, 1975). Winter (n.d.) suggests that attention is qualitatively different from other scarce resources and offers a set of satisficing rules for its allocation. His rules are more compatible with the empirical work presented in this paper than are the more formal economic analyses.

[10] Mintzberg (1970, 1973) provides the best published description of managerial observation methodology. The primary differences between his methodology and that used here is that in this study descriptions of all activities one minute or longer in duration were recorded. Mintzberg omitted all secretarial, personal, and social activities.

[11] The extent to which an observer's presence affected the managers and those around them, and therefore also affected the resulting data, can be partially inferred from four separate findings. One obvious measure is how much time each manager spent talking to me instead of doing whatever s/he would have done had I not been present. The average amount of time per day of observation spent by each manager talking to me was nine minutes; this represents 1.7% of the average day. A second indicator of my obtrusiveness is that in 300 hours of observation only five times was I asked to leave a manager's presence. On three occasions Manager A asked me to wait outside while he disciplined a student. Twice (on the same day) I was asked to leave an agency meeting in which a hiring decision was being debated. A third indication that the managers were not "putting on a show" for me was the amount of time they spent socializing and

conducting personal business. Had they wanted to impress me, they might have kept these activities to a minimum. In fact they constituted 14% (s.d.=7%) of all activity. Finally, at the end of each day of observation I asked the manager how I might improve my observation technique in later sessions: Had I been distracting? What should I do differently? No manager made any suggestion for improvement. They all indicated they did not find my presence troublesome and most of them said something similar to Manager A's comment: "I forgot you were there."

[12] See Sproull and Sproull (1981) for an extensive description of the general method of data management and analysis.

[13] Workday refers to the entire period from the time the manager arrives at the office in the morning (or at his/her first scheduled appointment, if before arriving at the office) to the time s/he leaves for home in the evening. Managers report that they also engage in about 45 minutes of work on average each evening, usually reading or paper work or occasionally attending a meeting.

[14] Program information includes program planning and monitoring, forecasting and scheduling the workload, and describing the program to the public. Logistics information includes that related to personnel deployment (e.g., attendance and recruiting) and materials and supplies requisition and distribution. External information includes that found in most incoming mail, journals, and requests and requirements from the larger organization. Interpersonal information includes that related to personnel counseling and social pleasantries. Examples of attention to each different category of

information may help sharpen the distinctions among them. Attention to program information: Manager A spent 5 minutes observing a class and afterwards talking to the teacher about how to improve her performance. Manager G spent 27 minutes describing to union and management officials how the first collective bargaining election would be conducted. Attention to logistics information: Manager A spent 9 minutes showing a carpenter where and how to put up some shelves in the program supply room. Manager G spend 7 minutes on the phone trying to buy a postage scale. Attention to external information: Manager A spent 45 minutes surveying students on whether or not they were bilingual because the larger organization required the information. Manager G spent 10 minutes filling out a survey on affirmative action in state agencies. Attention to interpersonal information: Manager A spent 14 minutes talking with students about how to find a summer job. Manager G spent 4 minutes talking with a staff attorney about the opening of baseball season. See Sproull (1977) for more detail on the categories of information.

15 Education organizations are characterized by an unclear technology (March and Olsen, 1976). One researcher suggested that this should lead to fewer explicit inspection activities on the part of managers and more "face work" or social maintenance activities (Weick, 1976: 13). The relatively greater attention paid to program core activities on the part of agency managers and the relatively greater attention paid to social pleasantries on the part of school managers support this hypothesis.

16 Thus the standard time allocation consulting advice always recommends that managers buffer themselves from their environment. See McKenzie (1972) for

one example.

17 The event unit allows the researcher to mark break points without having to worry about the importance or centrality of what is being interrupted. The activity unit allows the researcher to mark termination points without having to worry about whether or not related activities will follow. The topic unit allows the researcher to aggregate related activities, after the fact.

18 Recent work by March and his associates is an exception. See Cohen, March and Olsen (1972); March and Olsen (1976); Sproull, Weiner, and Wolf (1978) for analyses of how problems cycle in and out of attention over time.

REFERENCES

Ackoff, Russell. 1967. "Management misinformation systems." Management Science, 14: 145-156.

Allison, Graham. 1971. Essence of Decision. Boston: Little Brown.

Atkinson, John. 1969. "Change of activity." In T. Mischel (ed.) Human Action: 105-133. New York: Academic Press.

Clark, Burton. 1972. "The organizational saga in higher education." Administrative Science Quarterly, 17: 178-184.

Cohen, Michael and James March. 1974. Leadership and Ambiguity. New York: McGraw Hill.

Crecine, John P. 1969. Governmental Problem Solving. Chicago: Rand McNally.

Cyert, Richard and James March. 1963. A Behavioral Theory of the Firm. Englewood Cliffs, N.J.: Prentice-Hall.

Dubin, R. and S.L. Spray. 1964. "Executive behavior and interaction." Industrial Relations, 3: 99-108.

Freeland, James and Charles Stabell. 1978. "Allocation of managerial effort: An investigation of the relationship between decision strategies, environment, and performance." Behavioral Science, 23: 234 - 255.

Horne, J.H. and Tom Lupton. 1965. "The work activities of 'middle' managers--an exploratory study." The Journal of Management Studies, 2: 14-33.

Kahneman, Daniel. 1973. Attention and Effort. Englewood Cliffs, N.J.: Prentice-Hall.

Keegan, Warren. 1974. "Multinational scanning: A study of information sources utilized by headquarters executives in multinational companies." Administrative Science Quarterly, 19: 411-421.

Larkey, Patrick. 1979. Evaluating Public Programs: The Impact of Revenue Sharing on Municipal Governments. Princeton, N.J.: Princeton University Press.

Mackenzie, R. Alec. 1972. The Time Trap. New York: McGraw Hill.

March, James and Johan Olsen. 1976. Ambiguity and Choice in Organizations. Bergen: Universitetsforlaget.

McCall, Morgan, Ann Morrison, Robert Hannan. 1978. "Studies of managerial work: Results and methods." Greensboro, N.C.: Center for Creative leadership Technical Report Number 9.

Mintzberg, Henry. 1970. "Structured observation as a method to study managerial work." The Journal of Management Studies, 7: 87-104.

Mintzberg, Henry. 1973. The Nature of Managerial Work. New York: Harper and Row.

Mostofsky, David (ed.). Attention: Contempory Theory and Analysis. New York: Appleton-Century-Crofts.

Pounds, William. 1969. "The process of problem finding." Industrial Management Review, 11:1-19.

Radner, Roy. 1975. "A behavioral model of cost reduction." The Bell Journel of Economics, 6: 196-215.

Radner, Roy and Michael Rothchild. 1975. "On the allocation of effort." Journal of Economic Theory, 10: 358-376.

Robinson, John. 1977. How Americans Use Time: A Social-Psychological Analysis of Everyday Behavior. New York: Praeger.

Shlaim, Avi. 1976. "Failures in national intelligence estimates: The case of the Yom Kippur war." World Politics, 28: 348-380.

Simon, Herbert. 1953. "Birth of an organization: The Economic Cooperation Administration." Public Administration Review, 13: 227-236.

Simon, Herbert. 1971. "Designing organizations for an information - rich world." In Martin Greenberger (ed.). Computers, Communications and the Public Interest: 37-52. Baltimore: Johns Hopkins Press.

Simon, Herbert. 1973. "Applying information technology to organization design." Public Administration Review, 33: 268-278.

Snyder, Neil and William Glueck. 1980. "How managers plan--The analysis of managers' activities." Long Range Planning: 70-76.

Spence, Donald. 1973. "Analog and digital descriptions of behavior." American Psychologist, 23:479 - 488.

Sproull, Lee. 1977. Managerial attention in new education programs: A micro-behavioral study of program implementation. Stanford University: Unpublished doctoral dissertation.

Sproull, Lee and Robert Sproull. "Microethnography: A strategy for research in organizations." Carnegie-Mellon University: Unpublished paper.

Stewart, Rosemary. 1967. Managers and Their Jobs: A Study of the Similarities and Differences in the Way Managers Spend Their Time. London: McMillan.

Szalai, Alexander. 1966. "Trends in comparative time budget research." American Behavioral Scientist, 9: 3 - 8.

Szalai, Alexander, Phillip Converse, Pierre Feldheim, Erwin Scheuch, and Phillip Stone. 1972. The Use of Time. The Hague, Netherlands: Mouton.

Weick, Karl. 1976. "Educational organizations as loosely-coupled systems." Administrative Science Quarterly, 21: 1-19.

Winter, Sidney. n.d. "Notes on the economics of attention." New Haven: Yale University unpublished paper.

Wohlstetter, Roberta. 1963. Pearl Harbor. Stanford: Stanford University Press.

Wright, J. Patrick. 1979. On a Clear Day You Can See General Motors. New York: Avon.

Table 1

Descriptive Characteristics of Programs in Which Managerial Attention was Observed

Charactistics	Program					
Context	**A**	**B**	**C**	**D**	**E**	**F/G**
Location	Urban	Urban	Small City	Urban	Small City	State Capital
Larger Organization	School Dist.	Comm. College Dist.	School Dist.	School Dist.	School Dist.	State Government
Larger Org. Enrollment	~ 54,000	~ 10,000	~ 50,000	~38,000	~29,000	N.A.
Larger Org. Budget	~ $93million	~ $25million	~ $89million	~ $70million	~$37million	N.A.
Program						
Annual Budget	$650,000	$350,000	$200,000	$900,000	$160,000	$300,000
Professional Staff	9	10	5	12	5	2-9
Clerical Staff	1	3	1	3	1	3-7
Enrollment	250	300	80	260	75	N.A.
Type of Program	Career Educ.	Nontraditional	Career Educ.	Bilingual	Career Educ.	Labor relations
Age of Program when Observation began	3 mos.	1 mo.	3 mos.	3 mos.	3 mos.	-2 mos.
Manager						
Job Title	Principal	President	Program Co-ordinator	Director	Head teacher	Asst. to board/ Exec. director
Amt. previous managerial experience	1 yr.	2 yrs.	2 yrs.	none	1 yr.	2 yrs./none
Amt. previous experience in same context	3 yrs.	none	10 yrs.	2 yrs.	12 yrs.	2 yrs./none
Age	30-35	40-45	40-45	30-35	45-50	25-30/25-30
Sex	Male	Female	Male	Female	Male	Male/Male
Race	Black	White	White	Chicana	White	White/White

Table 2

Proportion of Conversation by Medium and Other Participant

Medium	Other Participant			
	Program Participants	Larger Organization	Public and Peers	Total*
Unscheduled Meetings	37.0%	3.9%	2.0%	42.9%
Scheduled Meetings	24	13	4.0	41
Telephone	4.8	3.7	6.3	14.8
Total	65.8	20.6	12.3	98.7

* Represents 80% of total day.

Table 3

Comparative Time Allocation Data for Different Kinds of Managers

TIME ALLOCATION	Chief Executives		Mid-Level Managers		Lower-Level Manager	
	Chief Executive Officers[a]	American College Presidents[b]	Middle Managers[c]	Middle Managers[d]	School and Agency Managers	Elementary Principal[e]
Medium (% of Day)						
Scheduled	35		34	10	34%	26
Unscheduled	32		26	55	34	40
Phone	6		6	9	13	10
Total Verbal Interaction	78	75	66	74	81	76
Alone	22	25	34	26	19	24
Participants (% of Interaction)						
Subordinates	48	52	37	--	66	70
Larger Organization	7	8	52	--	21	6
Peers and Public	45	40	11	--	12	24
Rhythm						
M-F Workweek	44h	50-55h	42h	--	43.5h	45h
% of Day at office (or "on campus")	62	50	75	85	74	--
% of Interaction with 1 other per.	48	47	32%	25%	60%	89
No. Verbal Contacts per day	15		--	--	45	

---

a. Mintzberg (1973)
b. Cohen & March (1974)
c. Stewart (1967)
d. Horne & Lupton (1965)
e. Wolcott (1973)

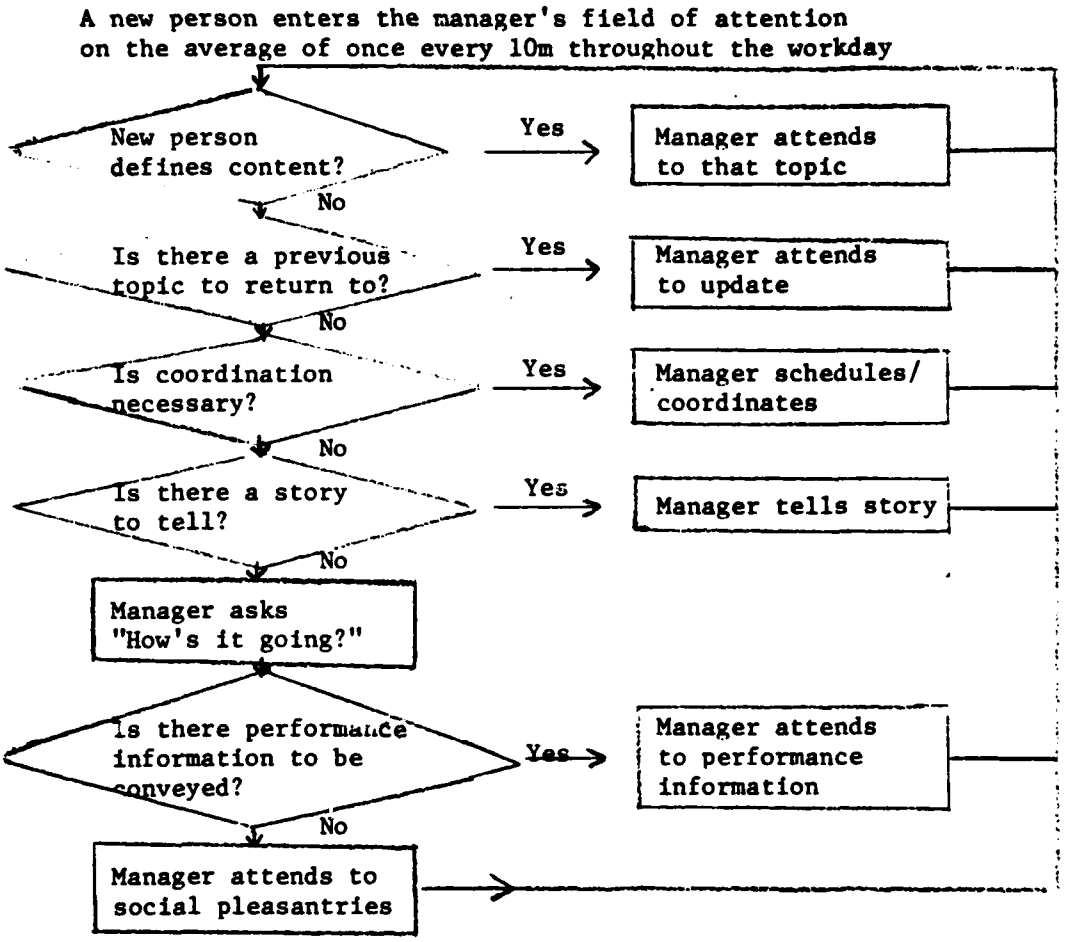
Table 4

Distribution of Attention to Subject Matter for School-Based and Agency-Based Managers

SUBJECT MATTER	% OF ATTENTION		
	School-Based Managers (n=5)	Agency-Based Managers (n=2)	Difference (1-2)
Program			
Technical Core	12.4	29.8	-17.4
Evaluation/Monitoring	6.7	3.0	3.7
Public Information	2.5	5.0	-2.5
Forecasts, Schedules & Records	2.3	7.5	-5.2
(Total)	(23.9)	(45.3)	(-21.4)
Logistics			
Personnel Deployment	10.8	14.6	-3.8
Materials & Supplies	6.4	23.1	-16.7
Desk	4.0	2.6	1.4
(Total)	(21.2)	(40.3)	(-19.1)
Environment		3.8	11.1
External Requirements	11.3	0.2	4.4
External Information	8.0	3.6	(15.5)
(Total)	(19.3)	(3.8)	
Interpersonal Relations			
Social Pleasantries	10.9	3.6	7.3
Personnel Counseling	2.6	0.8	1.8
(Total)	(13.5)	(4.4)	(9.1)
Review	14.2	2.8	11.4
Other			
Personal	5.8	2.8	3.0
Talk to Researcher	2.2	0.5	1.7
Unknown	0.2	--	.2
TOTAL	100.3	99.9	

Figure 1

Allocation Rules for the Topic of Conversation, Subject to Interrupt Rules




DAT
FILM